

AD _____
(Leave blank)

Award Number:
W81XWH-10-1-0151

TITLE:
Screening and Selection of New Antagonists of the RING-Mediated Hdm2/Hdmx Interaction

PRINCIPAL INVESTIGATOR:
Professor Julio A. Camarero, PhD

CONTRACTING ORGANIZATION:
University of Southern California

Los Angeles, CA 90033

REPORT DATE:
May 2013

TYPE OF REPORT:
Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

*Form Approved
OMB No. 0704-0188*

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE Tæ Å 2013			2. REPORT TYPE Final		3. DATES COVERED 1 Tæ & @ 2010 – G OI 12 2013	
4. TITLE AND SUBTITLE Screening and Selection of New Antagonists of the RING-Mediated Hdm2/Hdmx Interaction			5a. CONTRACT NUMBER			
			5b. GRANT NUMBER W81XWH-10-1-0151			
			5c. PROGRAM ELEMENT NUMBER			
6. AUTHOR(S) Julio A. Camarero, Krishnapa Jagadish, Yanbin Ji Mark Wade, Geoffrey Wahl E-Mail:			5d. PROJECT NUMBER			
			5e. TASK NUMBER			
			5f. WORK UNIT NUMBER			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Salk Institute Univ. of Southern California Road La Jolla, CA 92037 Los Angeles, CA 90033			8. PERFORMING ORGANIZATION REPORT NUMBER			
			10. SPONSOR/MONITOR'S ACRONYM(S)			
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012			11. SPONSOR/MONITOR'S REPORT NUMBER(S)			
			12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for Public Release; Distribution Unlimited			
13. SUPPLEMENTARY NOTES						
14. ABSTRACT Prostate cancer poses a major public health problem in the United States and worldwide. It has the highest incidence and is the second most common cause of cancer deaths in North American men resulting in over 30,000 deaths per annum. Consequently, there is an urgent need to develop novel therapeutic approaches. We propose to use a novel cyclotide-based molecular scaffold for generating molecular libraries that will be screened and selected in vivo to identify antagonists of the RING-mediated Hdm2/Hdmx interaction. Our innovative approach will use cell-based <i>E. coli</i> libraries in which individual bacteria express a different cyclotide. This comprises a new single cell-single compound approach to identify protein-protein binding antagonists. These compounds will be subjected to a two-step screen, the first involving a high throughput FRET-based FACS screen in bacteria, and the second a bioluminescence complementation assay in cancer cells. Together, these assays will identify cyclotides that disrupt Hdm2-Hdmx interactions, activate p53, and elicit p53-dependent cytotoxicity in prostate cancer cells. During the second year of this project we have accomplished the following: 1) Developed a more efficient method for the biosynthesis of cyclotides in <i>E. coli</i> cells; 2) We have validated our FRET-based reporter to sort cell populations; 3) Developed and characterized a novel cyclotide with p53 activating properties.						
15. SUBJECT TERMS						
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON USAMRMC	
a. REPORT U	b. ABSTRACT U	c. THIS PAGE U	UU	120	19b. TELEPHONE NUMBER (include area code)	

	<u>Page</u>
Introduction.....	4
Body.....	5
Key Research Accomplishments.....	16
Reportable Outcomes.....	16
Conclusion.....	19
References.....	21
Abbreviations.....	22
Appendices.....	23

Introduction

Background

Prostate cancer poses a major public health problem in the United States and worldwide. It has the highest incidence and is the second most common cause of cancer deaths in North American men resulting in over 30,000 deaths per annum. Consequently, there is an urgent need to develop novel therapeutic approaches. The molecular mechanisms of development and progression of prostate cancer are complicated and likely to involve multiple factors. The human double minute 2 (Hdm2 or human Mdm2) protein is amplified or overexpressed in a number of human tumors, including prostate cancer. Importantly, the Hdm2 antagonist nutlin-3, which is particularly effective in causing p53-dependent apoptosis in Hdm2-amplified cultured cells, exhibits antitumor activity on human prostate LNCaP and other xeno-grafts in nude mice. Hdm2 promotes p53 degradation through an ubiquitin-dependent pathway (Fig. 1). The exact mechanism by which p53 is stabilized is unclear, although a series of post-translational modifications to itself, Hdm2 and the closely related protein Hdmx (or human MdmX, which is also known as Hdm4 or human Mdm4), are thought to dissociate the p53-Hdm2 complex leading to increased levels of p53. Although Hdmx contains a RING domain that is very similar to the RING domain of Hdm2, it does not possess intrinsic E3 ubiquitin ligase activity. Dimerization, mediated by the conserved C-terminal RING domains of both Hdm2 and Hdmx, appears to greatly augment this activity. While the Hdm RING domains can form homodimers, heterodimers form preferentially resulting in reduced auto-ubiquitylation of Hdm2 and increased p53 ubiquitylation. Thus disruption of this interaction should inactivate Hdm2 E3 ligase activity and consequently increase p53 abundance. The recent elucidation of the structure of the complex formed by the RING domains of Hdm2 and Hdmx suggests the feasibility of obtaining Hdm-specific E3 ligase inhibitors by targeting the Hdm2/HdmX RING domain dimer interface rather than the primary E2 binding site that is common to many RING domain E3-ubiquitin ligases (Fig. 2).

Objectives

Disruption of Hdm2 function is a very novel attractive therapeutic target for prostate cancer [3, 4] (Fig. 1). As mentioned earlier, nutlin-3, a peptidomimetic that disrupts the p53-Hdm2 interaction, activates p53 pathways both *in vitro* and *in vivo* in human cell lines that possess wild-type p53 and overexpress Hdm2 [5]. Vousden and colleagues have also established that it is possible to stabilize p53 by directly inhibiting the E3-ligase activity of Hdm2 [6], although this approach may also inhibit other E3s. The recent elucidation of the structure of the Hdm2-Hdmx RING domain heterodimer shows that both protein domains contribute residues for E3-ligase activity (Fig. 3). This strongly suggests that it might be possible to obtain Hdm-specific E3 ligase

Figure 1. Hdm2-Hdmx may be functionally inhibited at multiple steps to reactivate p53 function. Numbered circles indicate potential therapeutic targets for the development of Hdm2/Hdmx antagonists. In this proposal will target step 3 highlighted in green.

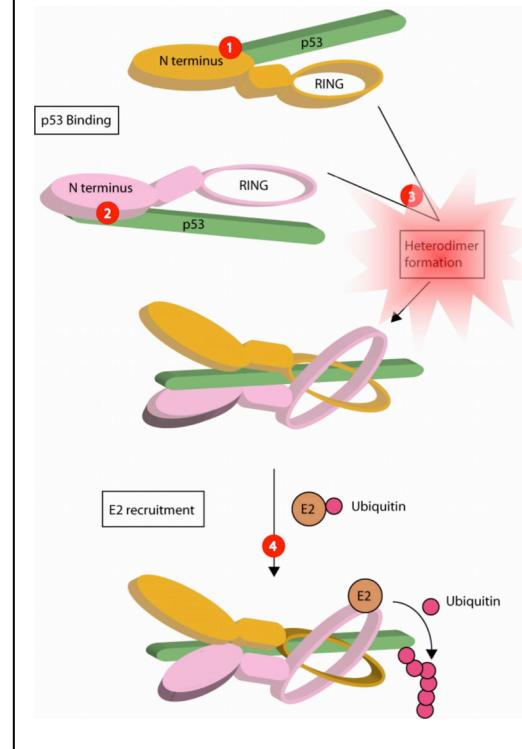
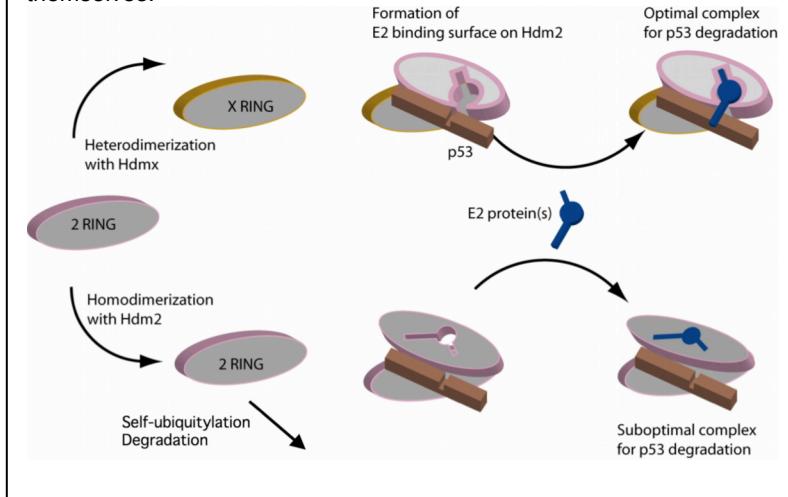


Figure 2. Hdm2-Hdmx heterodimers are more effective p53 ubiquitin ligases than Hdm2 homodimers, which can also self-ubiquitylate themselves.



inhibitors by targeting the Hdm2-HdmX RING domain dimer interface rather than the primary E2 binding site that is common to many RING domain E3-ubiquitin ligases.

To achieve this objective we are using cell-based libraries of cyclotides for selecting specific cyclotide sequences able to antagonize the RING-mediated interaction between Hdm2 and Hdmx to obtain Hdm-specific E3 ligase inhibitors. The use of the modified protein splicing technology developed in the Camarero lab allows to generate large, genetically-encoded cyclotide libraries in bacterial cells. These cell-based libraries are then screened using an in-cell FRET-based reporter in combination with high throughput flow cytometry to identify bacteria encoding cyclotides able to disrupt Hdm2-Hdmx interactions. Selected cyclotides are then characterized by NMR and assayed in mammalian cells using the BiLC assay implemented in p53 wild-type and p53-null cancer cell lines to ascertain p53-dependent biological activity.

This proposal represents a novel approach for antagonizing Hdm2-Hdmx E3-ligase activity. Selected cyclotides will be highly specific for antagonizing Hdm2 E3-ligase, and for eliciting p53-dependent cytotoxicity in cancer cells. The use of the cyclotide scaffold will enable these peptide antagonists to have the required increased stability, cellular membrane penetration, proteolysis resistance and serum clearance needed to be considered as viable drug development candidates. It is also important to remark, that this cell-based technology could be easily adapted to screen for antagonists for other relevant protein-protein interactions in prostate cancer; for example those that may induce tumor cell apoptosis independent of p53 or compounds able to reactivate mutant p53. These compounds could be used in combination with Hdm2 antagonists to prevent tumor relapse or secondary tumor formation.

Summary of results found for FY2011-2012

During the second year of this project we have accomplished the following: 1) Developed a more efficient method for the biosynthesis of cyclotides in *E. coli* cells, which allows the incorporation of non-natural amino acids using nonsense tRNA suppressor technology; 2) We have developed and validated a FRET-based reporter to sort cell populations expressing interacting and non-interacting RING-domains and 3) Developed and characterized a novel cyclotide with p53 activation properties and trigger cytotoxicity in p53-wt prostate cancer cells and xenograft animal models.

Body

A. Specific Aims

We are using a cyclotide-based molecular scaffold for generating molecular libraries that are screened and selected *in vivo* for potential antagonists for the RING-mediated Hdm2/Hdmx interaction. In this innovative approach, we are using cell-based libraries (*E. coli* cell libraries) where every single cell will express a different cyclotide, in what we could call a single cell-single compound approach. These compounds are then screened and selected for their ability to inhibit the Hdm2/Hdmx interaction inside the bacterial cell using a genetically-encoded FRET-based reporter [7] in combination with high throughput flow cytometry to identify bacteria encoding cyclotides able to disrupt Hdm2-Hdmx interactions. This screening assay is optimized to be used in *E. coli* in combination fluorescence activated cell sorting (FACS) and is designed to minimize the number of false positives. The Camarero lab has developed a similar assay to screen cyclotides able to inhibit *B. anthracis* Lethal Factor protease activity anthrax toxin binding, demonstrating the feasibility of this approach for use in bacteria.

Selected cyclotides are also characterized by NMR and assayed in mammalian cells using the BiLC assay implemented in p53 wild-type and p53-null cancer cell lines to ascertain p53-dependent biological activity. The BiLC assay developed in the Wahl lab shows a high dynamic range, high degree of reproducibility and will be used to validate the ability of any cyclotide selected in bacteria to enter mammalian cells in concentration sufficient to antagonize Hdm2-Hdmx RING-mediated interaction.

Specific Aim 1. To screen and select cyclotide-based peptides able to disrupt the Hdm2-Hdmx RING heterodimer. The objectives of this aim are the production of large genetically-encoded libraries of cyclotides in living *E. coli* cells ($\approx 10^9$) and the development of FRET-based *in vivo* screening reporter to select cyclotides able to inhibit Hdm2-Hdmx RING heterodimer. Cells able to express active cyclotides will be selected using high throughput flow cytometry methods such as fluorescence activated cell sorting (FACS)

Specific Aim 2. To test and evaluate the inhibitory and biological activity of selected cyclotides. Selected cyclotides will be tested *in vitro* using a combination of fluorescence assays and nuclear magnetic resonance

(NMR). Biological activity will be assayed using different cancer cell lines to evaluate their ability to activate endogenous p53.

B. Studies and results

1) Biosynthesis and characterization of genetically-encoded cyclotide-based libraries.

Our group has recently developed and successfully used a biomimetic approach for the biosynthesis of folded cyclotides (Fig. 3) inside cells by making use of modified protein splicing unit in combination with expressed protein ligation (EPL) [8]. Using this approach, we have biosynthesized a small genetically-encoded library based on the cyclotide MCoTI-I [9]. In the second year of this project we have developed a new approach for the efficient production of cyclotides in bacterial cells using protein trans-splicing (PTS) (Fig. 4) (Appendix: paper #7). Using this new approach we have shown that we can express folded cyclotides inside living

cells 10 times more efficiently than with the previous method using intracellular EPL [8, 10]. Importantly, the higher efficiency of PTS-mediated cyclization combined with nonsense-codon suppressor tRNA technology has also made possible for the first time the in-cell production of cyclotides containing unnatural amino acids (UAAAs). These results open the exciting possibility for the generation of genetically-encoded cyclotide-based libraries containing additional chemical diversity for selection on novel cyclotide-based sequences with improved biological activity. For example the inclusion of amino acids with side-chains containing classical warheads as ketoamides, boronates and hydroxamates, which are absent from the 20 proteinogenic amino acids, should help in screening and selection of cyclotides with more effective pharmacological properties.

Protein trans-splicing is a post-translational modification similar to protein splicing with the difference that the intein self-processing domain is split into two fragments N- (I_N) and C-intein (I_C). The split-intein fragments are not active individually, however, they can bind to each other with high specificity under appropriate conditions to form an active protein splicing or intein domain in trans.[2] PTS-mediated backbone can be accomplished by rearranging the order of the intein fragments, i.e. by fusing the I_N and I_C fragments to the C- and N-terminus of the polypeptide to cyclize, the trans-splicing reaction results in the formation of a backbone-cyclized polypeptide (Fig. 3). This approach has been recently used for the biosynthesis of small cyclic hexapeptides [11]. In this work in-cell cyclization was performed using the naturally occurring *Synechocystis* sp. (Ssp) PCC6803 DnaE split intein [12]. The use of Ssp DnaE intein, however, requires the presence of specific amino acid residues at both intein-extein junctions for efficient protein splicing [13]. To overcome this problem we used the *Nostoc punctiforme* PCC73102 (Npu) DnaE

Figure. 3. Primary and tertiary structure of cyclotides from the plants *Momordica cochinchinensis* (MCoTI-II) and *Oldenlandia affinis* (Kalata B1).

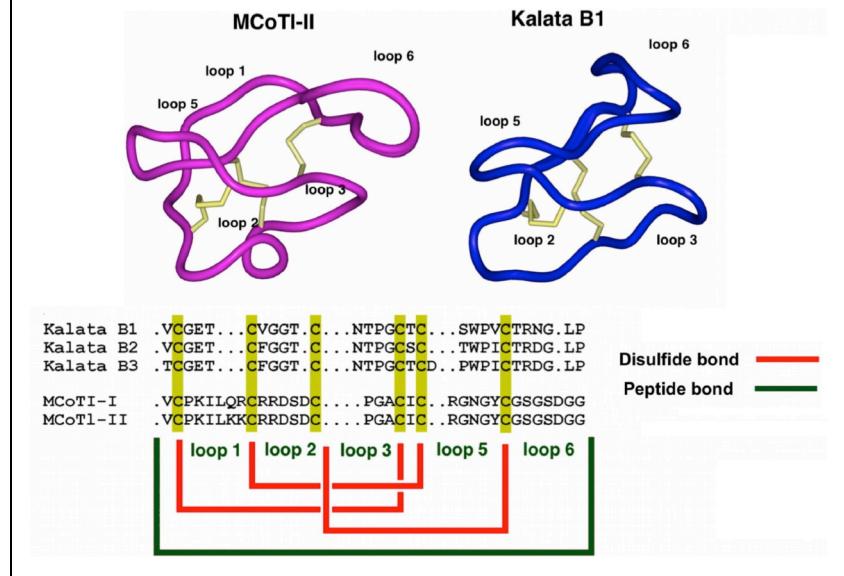
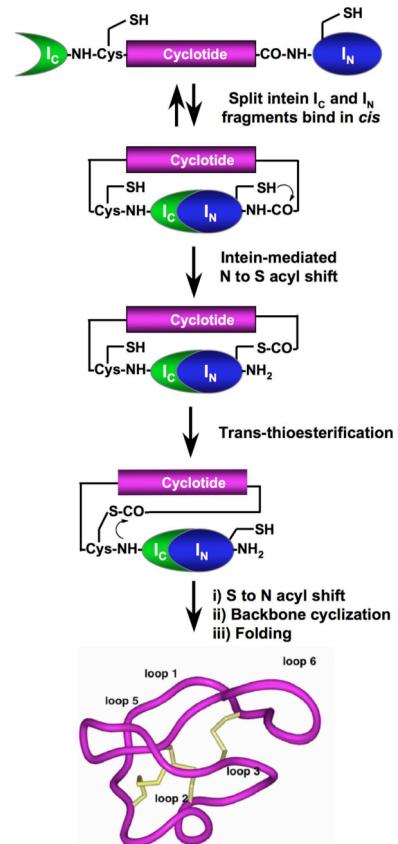


Figure. 4. Scheme depicting the in-cell biosynthesis of cyclotides using protein trans-splicing (PTS)-mediated backbone cyclization.

PTS-mediated backbone cyclization of cyclotides

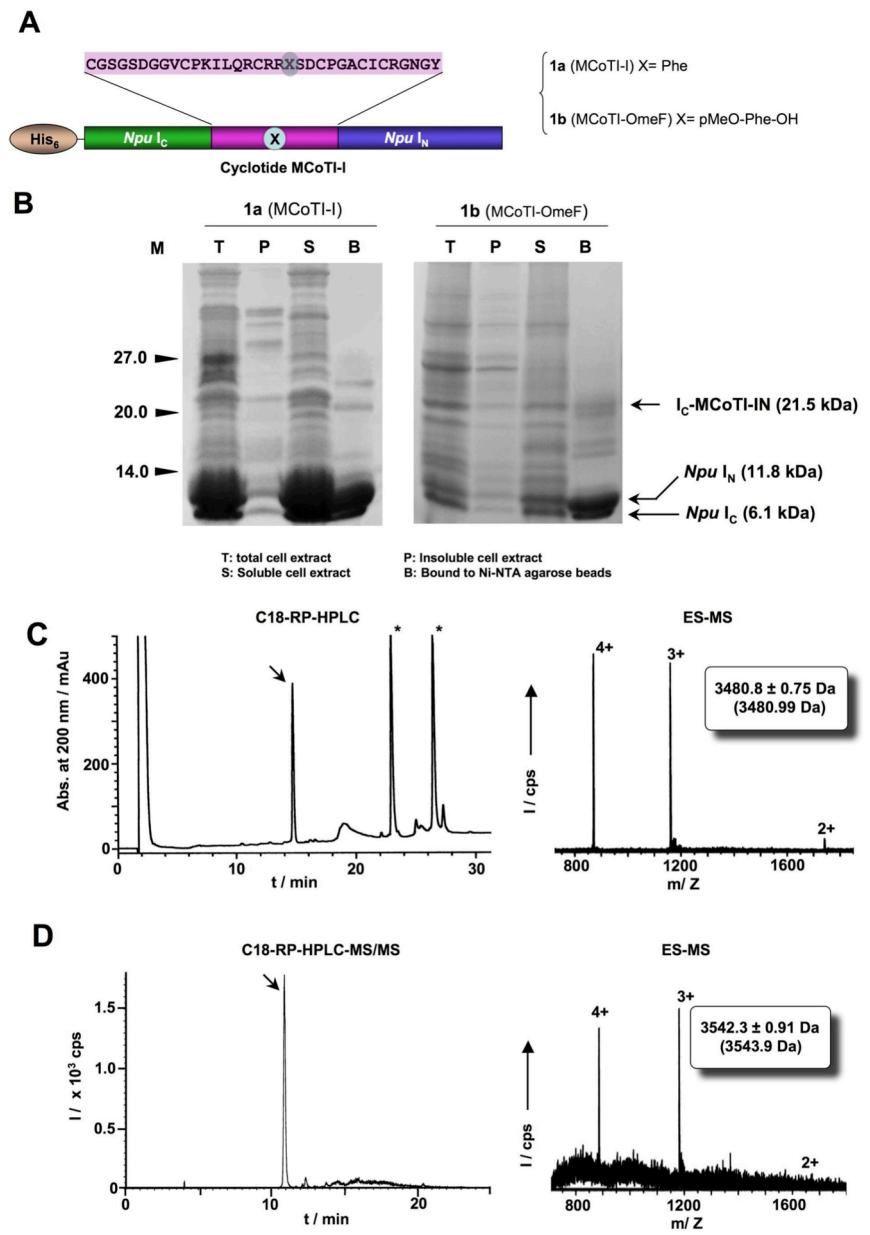


split-intein instead. This DnaE intein has been reported to have the highest rate of protein trans-splicing ($\tau_{1/2} \approx 60$ s) [14] and also high splicing yield [14, 15].

We explored first the ability of the *Npu* DnaE split-intein to produce folded wild-type MCoTI-I cyclotide inside living *E. coli* cells. MCoTI-I is a powerful trypsin inhibitor recently isolated from the seeds of *Momordica cochinchinensis*, a plant member of cucurbitaceae family. To accomplish this, we designed the split-intein construct 1a (Fig. 3). In this construct, the MCoTI-I linear precursor was fused in frame at the C and N termini directly to the *Npu* DnaE I_N and I_C polypeptides. No additional C- or N-extein native residues were added in this construct. We used the native Cys residue located at the beginning of loop 6 of MCoTI-I (Figs. 3 and 5) to facilitate backbone cyclization. A His-tag was also added at the N-terminus of the construct to facilitate purification.

In-cell expression of wild-type MCoTI-I using PTS-mediated backbone cyclization was achieved by transforming the plasmid encoding the split-precursor 1a into Origami(DE3) cells to facilitate folding, and then over-expressing the MCoTI-precursor split-intein for 18 h at room temperature. Using these conditions the precursor was expressed at very high levels (≈ 70 mg/L) and almost completely cleaved ($\geq 95\%$ in vivo cleavage, Fig. 5B). Reducing the induction time during the expression of precursor 1a did not decrease the level in vivo cleavage significantly, indicating the inherent ability of the construct to undergo protein trans-splicing. The high reactivity of this precursor prevented us to perform a full characterization of the precursor protein including kinetic studies of the trans-splicing induced reaction in vitro. Next, we tried to isolate the natively folded MCoTI-I generated in-vivo by incubating the soluble fraction of a fresh cell lysate with trypsin-immobilized agarose beads. As indicated before, MCoTI-cyclotides can bind to trypsin beads with high affinity only if they are correctly folded. Therefore this step can be used for affinity purification but also to test the biological activity of recombinantly produced cyclotides. After extensive washing, the absorbed products were eluted with a solution of 8 M guanidium chloride (GdmCl) and analyzed by HPLC. The HPLC analysis revealed the presence of a major peak that had the expected mass for the natively folded MCoTI-I (Fig. 5C). Recombinant MCoTI-I produced by PTS-mediated cyclization was also characterized by 2D-NMR spectroscopy and shown to be identical to natively folded MCoTI-I [10].

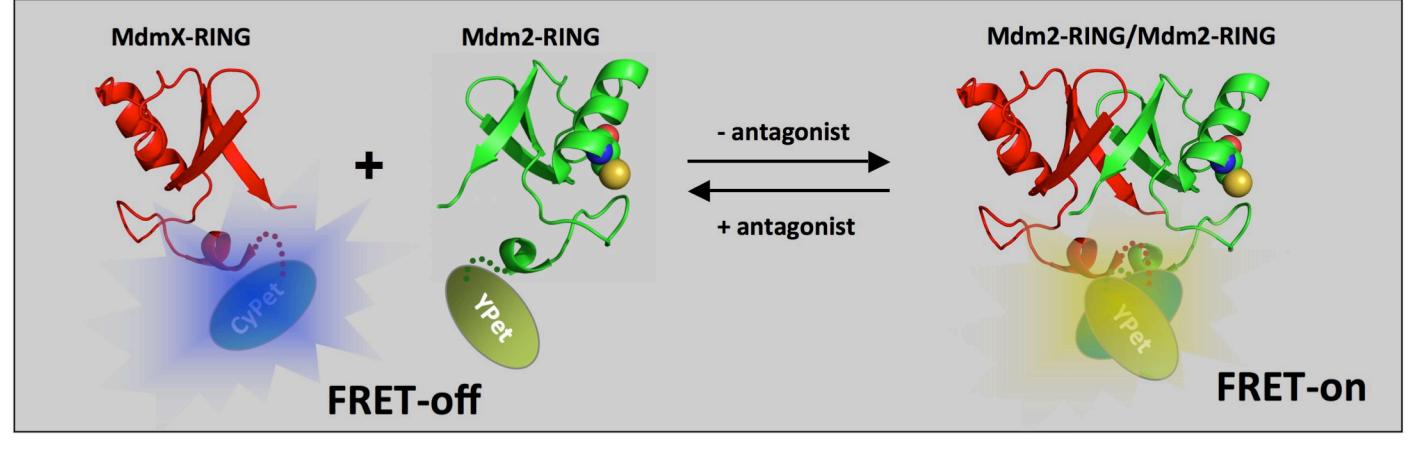
Figure 5. In-cell expression of MCoTI-I and MCoTI-OmeF (containing the unnatural amino acid p-MeO-Phe) in *E. coli* cells using PTS-mediated backbone cyclization. **A.** Design of the protein precursor for in-cell production of cyclotides MCoTI-I and MCoTI-OmeF. **B.** SDS-PAGE analysis of expression of precursors **1a** (MCoTI-I) and **1b** (MCoTI-OmeF) in *E. coli* Origami(DE3) cells. **C.** Analytical HPLC analysis of the soluble cell extract of bacterial cells expressing precursor **1a** (MCoTI-I) after pulldown with trypsin-agarose beads. Folded MCoTI-I is marked with an arrow. Endogenous bacterial proteins that bind trypsin are marked with an asterisk. **D.** Analytical HPLC-MS/MS analysis of the soluble cell extract of bacterial cells expressing precursor **1b** (MCoTI-OmeF) after pulldown with trypsin-agarose beads. Folded cyclotide is marked with an arrow.



The in-cell expression level of folded MCoTI-I produced by PTS-mediated cyclization was estimated to be ≈ 70 $\mu\text{g/L}$ of bacterial culture, which corresponds to an intracellular concentration of ≈ 7 μM . In-cell expression of folded MCoTI-cyclotides by PTS was about 10 times more efficient than intramolecular EPL-mediated backbone cyclization. This improvement could be explained first, by the choice of the split-intein (*Npu* DnaE) used, which is extremely efficient in combining fast kinetics and good yields of trans-splicing, and also by the differences in the cyclization process between the PTS and EPL methods. In PTS, the cyclization is driven by the affinity between the two-intein fragments, I_N and I_C , which in the case of the *Npu* DnaE intein is very tight ($K_D \approx 3$ nM) [16]. Once the intein complex is formed the trans-splicing reaction is also extremely fast ($\tau_{1/2} \approx 60$ s for the *Npu* DnaE intein). In contrast, EPL-mediated cyclization follows a slightly more complex mechanism that relies first in the formation of the C-terminal thioester at the N-extein-junction and removal of the N-terminal leading sequence (a Met residue in this case) to provide a N-terminal Cys. These two groups then react to form a peptide bond between the N- and C-termini of the polypeptide. It is also worth noting that in contrast with the *Ssp* DnaE intein, which requires at least 4 native residues at the N- and C-terminal extein-intein junctions to work efficiently [17], the *Npu* ortholog shows a good sequence tolerance for both junctions as is demonstrated by the efficient trans-splicing of precursor **1a** (Fig. 5A). The tetrapeptide sequences at both intein-extein junctions in construct **1a** have only a 20% sequence homology with the native sequences of both *Npu* DnaE exteins.

Encouraged by these results, we decided to try in-cell expression of cyclotide MCoTI-I incorporating UUAs using PTS. This cyclotide contains the UAA p-MeO-Phe-OH (OmeF) replacing residue Asp¹² in loop 2. Encoding of the UUA was carried out by replacing the Asp¹² residue in MCoTI-I by an amber stop codon (TAG). The Asp¹² residue is located in the middle of loop 2 (Fig. 3), which has been shown to be tolerant to mutations without affecting the structure and biological activity of the resulting cyclotide [10]. The incorporation of UUAs into the cyclotide framework was tested with *p*-methoxyphenylalanine (OmeF), which has been already successfully encoded into different recombinant proteins [18]. To accomplish this, precursor **1b** (Fig. 5A) was overexpressed in Origami(DE3), co-transformed with UUA-encoding plasmid pVLOmeR (a kind gift from Dr. Lei Wang, The Salk Institute), in the presence of 1 mM OmeF. Construct **1b** is similar to **1a** but was designed to incorporate UUAs into residue Asp¹² in MCoTI-I (Figs. 3 and 5A). The expression level of the intein precursors was ≈ 10 mg/L (15% suppression). In-vivo trans-splicing for **1b** was also similar ($\geq 90\%$, Fig. 5B) to

Figure 6. FRET-based reporter for screening Hdm2/MdmX RING-mediated Interactions. A. Principle for the FRET-based reporter to screen for antagonists against the RING-mediated Hdm2/MdmX heteronuclear complex. The formation of the complex brings in close proximity fluorescent proteins YPet and CyPet. Excitation of CyPet with blue light allows the transfer of energy to the yellow fluorescent protein YPet producing yellow fluorescence. Inhibition of the Hdm2/MdmX interaction prevents the transfer of energy from the blue fluorescent protein CyPet thus giving only blue fluorescence.



that of the wild-type PTS construct (**1a**). Cyclotide MCoTI-OmeF was purified by affinity chromatography using trypsin-agarose beads from fresh soluble cell lysates. The trypsin-bound fraction was analyzed by HPLC and LC-MS/MS (Fig. 5D). In-cell expression level for cyclotide MCoTI-OmeF was estimated to be ≈ 1 $\mu\text{g/L}$, which equals to an intramolecular concentration of 0.1 μM .

Figure 7. Genomic characterization of the MCoTI-library where loop 2 was completely randomized using an NNK scheme. A. DNA sequences of loop 2 for 60 different clones randomly chosen from a MCoTI-I loop 2 based library using the pTXB1 plasmid. B. Graphical representation of the amino acid composition of the 60 clones sequenced. The graph was produced using WebLogo 3 (<http://weblogo.threeplusone.com/>).

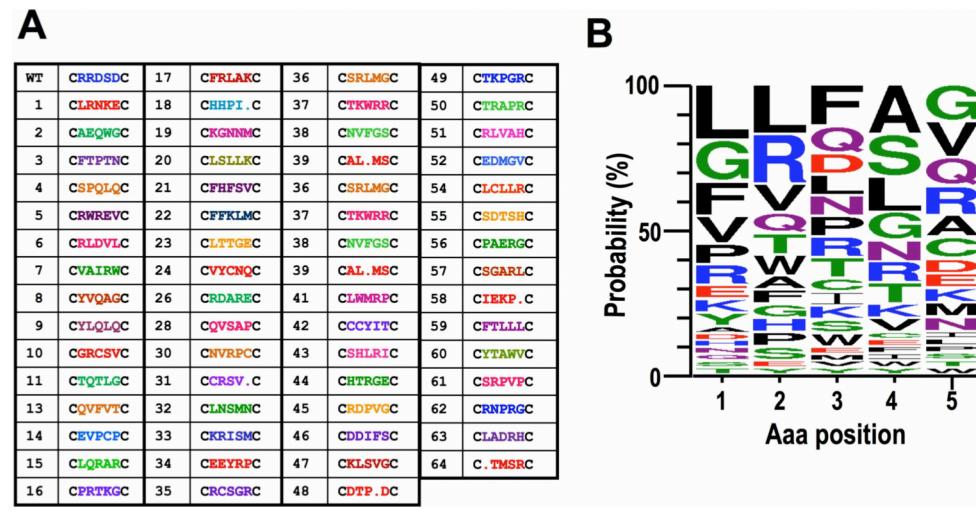
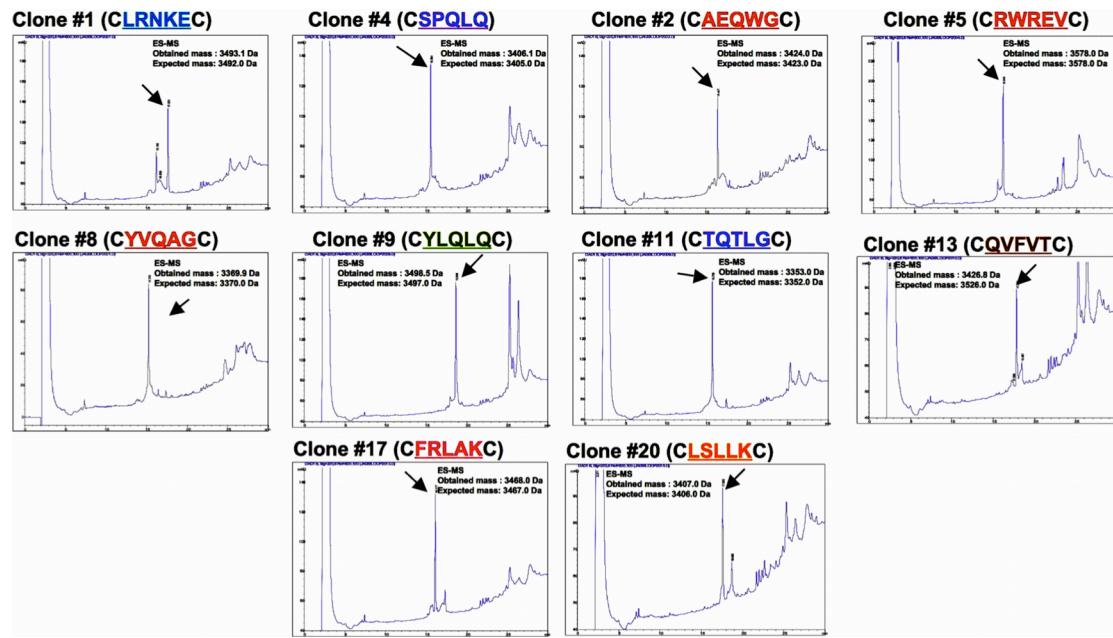


Figure 8. HPLC analytical traces of the cyclization/folding crudes for individual clones isolated from the MCoTI-I loop 2 library expressed using plasmid pTXB1 in *E. coli* before and after purification using trypsin-agarose beads. Folded cyclotides were correctly characterized by ES-MS (mass spectrometry) and their ability to bind trypsin.



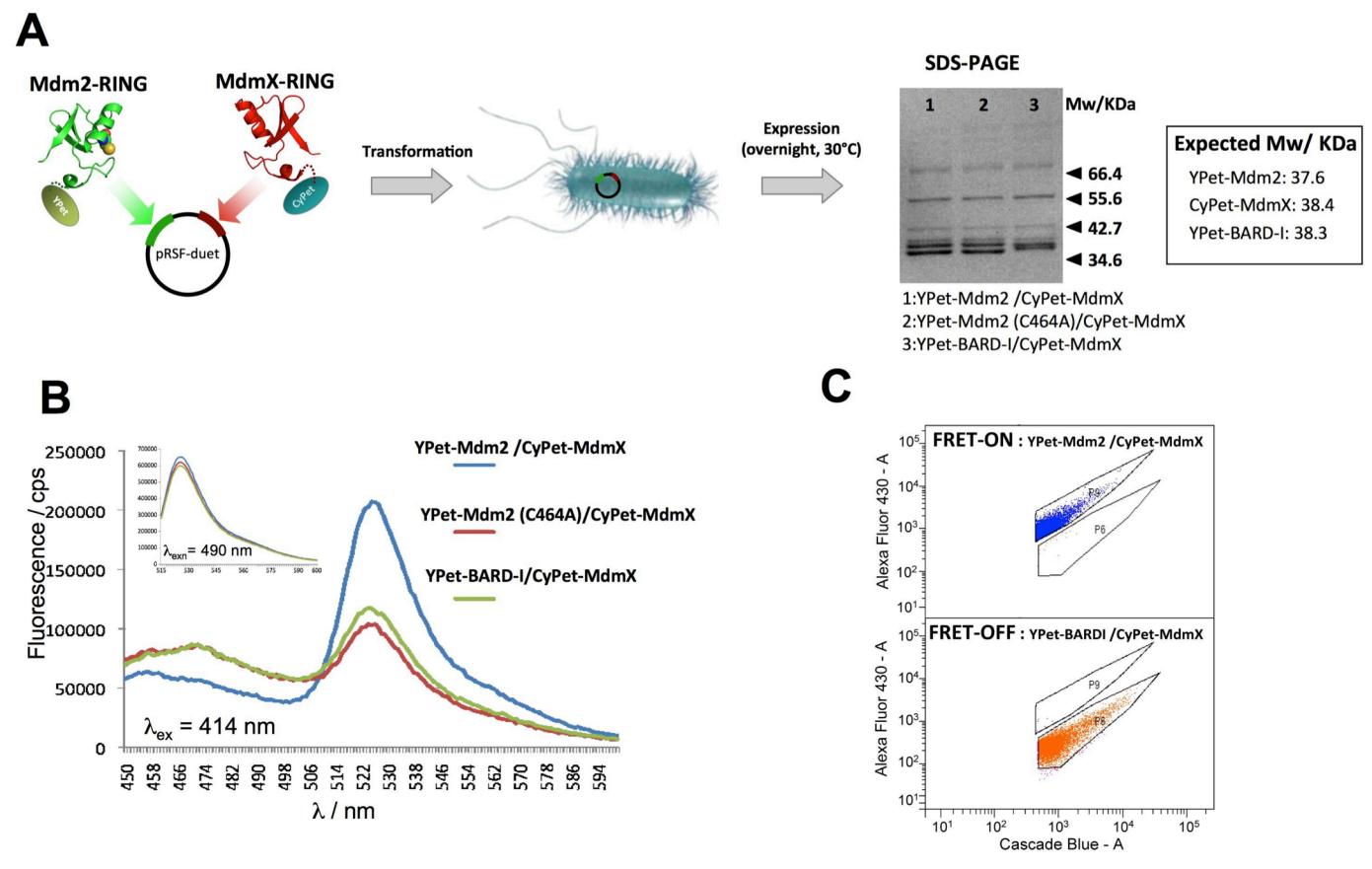
made possible for the first time the in-cell production of cyclotides containing UUAs. These results open the exciting possibility for the generation of genetically-encoded cyclotide-based libraries containing additional chemical diversity for the selection on novel cyclotide-based sequences with improved biological activity to target protein-protein interactions involved in the p53 pathway. We are currently using this approach for the production of cyclotide-based libraries using UAAs to target the RING-mediated interaction between Hdm2 and HdmX proteins.

We have also used PTS for the generation of libraries using the loop 2 of cyclotide MCoTI-I. This library had a diversity of 1.5 million different sequences (Figs. 7 and 8). These libraries are currently being screened for antagonists for the interaction between Hdm2 and HdmX.

In summary, we have shown that the biosynthesis of cyclotides can be achieved in live cells more efficiently with PTS than EPL. Importantly, in-cell PTS-mediated cyclization also allows the biosynthesis of cyclotides containing UAAs. Although PTS has been used earlier for the generation of small cyclic peptides using the less efficient *Ssp* DnaE split-intein [13], to our knowledge this is the first time that this method is used for the production of native folded cyclotides. We estimated that in-cell production of cyclotide MCoTI-I was around 10-times more efficient using *Npu* DnaE PTS than EPL, and therefore provides an attractive alternative for the recombinant production of these type of polypeptides. Importantly, the higher efficiency of PTS-mediated cyclization combined with nonsense-codon suppressor tRNA technology has also

2) *Cell-based reporter to screen in-cell RING-mediated Hdm2/HdmX interactions.* During the first year of the project we have generated a fluorescent reporter to screen RING-Hdm2/RING-HdmX interaction. The principle for this approach is depicted in Figure 6. Our FRET-based reporter system uses a CyPet and YPet fluorescent proteins, which are fused the N-terminus of the RING domains of Hdm2 (429-491 aa) and HdmX (427-490 aa), respectively. The N-terminal region of the Hdm2/X RING domains can easily tolerate the addition of different protein domains or protein fragments without altering their heterodimerization and biological function. For example the Wahl group has shown that half-luciferase fragments can be added without affecting the ability of

Figure 9. In-cell FRET-based reporter to screen inhibitors against RING-mediated Hdm2/HdmX interaction. **A.** Cloning and co-expression of YPet-Hdm2 and CyPet-HdmX fluorescent constructs on *E. coli*. YPet-Hdm2 and CyPet-HdmX were cloned into a pRSF duet expression vector. The resulting vector was used to transform *E. coli* cells and both proteins were expressed at room temperature for 18 h. Purified proteins were analyzed by SDS-PAGE. **B.** Fluorescence spectra of live *E. coli* cells expressing YPet-Hdm2 / CyPet-HdmX (FRET-on), YPet-Hdm2 (C464A) / CyPet-HdmX (FRET-off), and YPet-BARD-I / CyPet-HdmX (FRET-off). Excitation to quantify FRET signal was performed at 414 nm. Inset: quantification of fluorescent protein YPet in live *E. coli* cells expressing FRET-reporter indicates that the differences observed in FRET signal are not due to different expression levels of the YPet fusion protein. Cells were excited at 490 nm for YPet quantification. **C.** Analysis by FACS of live *E. coli* cells expressing YPet-Hdm2 / CyPet-HdmX (FRET-on); and YPet-Hdm2 (C464A) / CyPet-HdmX (FRET-off). The RING domains of Hdm2 (C464A) and BARD-I (26-126) do not interact with the RING domain of HdmX and were used as negative controls to evaluate the background fluorescence of the FRET-off state.



Hdm2 and Hdmx RING domains to heterodimerize. Moreover, to prevent any potential steric hindrance that could interfere with the molecular recognition process, we initially introduced the flexible linker [Gly-Gly-Ser]₅ between the interacting proteins or protein domains and the corresponding fluorescent proteins.

We have shown that this system can be used to monitor the RING-mediated interaction between Hdm2 and HdmX. As shown in Figure 9, when the Hdm2-HdmX RING heterodimer is formed in vivo the pair CyPEt-YPet exhibits high FRET signal indicating the formation of the complex. Background FRET signal was evaluated using cells expressing YPet-HdmX and the CyPet-Hdm2 mutant (C464A). This mutation in the RING domain of Hdm2 prevents Hdm2-HdmX heterodimerization and provides the background FRET signal when the Hdm2/HdmX complex is not formed. We

also tested the FRET background using a cell line co-expressing the RING domain of BARD-I (26-126) fused to YPet at its N-terminus with CyPET-HdmX as these two RING domains do not interact with each other. In both cases the background FRET signal when the RING heterodimerization was prevented (i.e. FRET-off state) was ≈ 2 times smaller than that of the positive control using CyPet-HdmX/YPet-Hdm2 or CyPet-Hdm2/YPet-HdmX (i.e. FRET-on state) (Fig. 9B). Analysis by fluorescence activated cell sorting (FACS) also revealed that both populations of cells, i.e.

FRET-on and FRET-off cells can be easily separated by FACS (Fig. 9C).

During the second year of the project we have developed a protocol for FACS to separate mixtures of cells depending on their FRET-state when using the fluorescence reporter described above. As shown in Figure 10, this protocol (summarized in Fig. 11) can be easily used to enrich a population of *E. coli* cells from an initially predominant FRET-ON phenotype (RING-domain complex formed) into a FRET-OFF phenotype (RING domain antagonized). In a model experiment performed with an artificial mixture of *E. coli* cells containing cells transfected with pRSF duet plasmids encoding

CyPet-HdmX / YPet-Hdm2 (FRET-ON population) and CyPet-Hdm2 / YPet-BARD-I (FRET-OFF population) was enriched from a population with a ratio FRET-OFF/FRET-ON of 1:100 to approximately 100:1 in just four rounds of sorting (Fig. 10). The enrichment process was also monitored by PCR and DNA sequencing as shown at the bottom of Figure 10.

Figure 10. Fluorescent-activated cell sorting enrichment of an artificial mixture of *E. coli* cells expressing CyPet-HdmX / YPet-Hdm2 (FRET-ON) and CyPet-HdmX / YPet-BARD-I (FRET-OFF) in a ratio FRET-ON/OFF of 100:1 to a final ratio FRET-ON/OFF of 4:96. Every enrichment step was monitored by DNA sequencing using a T7-terminator primer and by PCR using primers specific for the RING domains of Hdm2 and BARD-I.

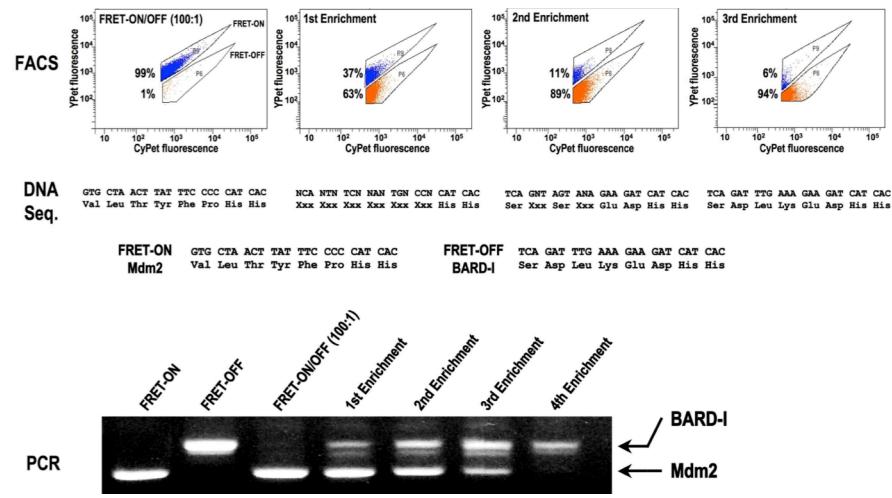
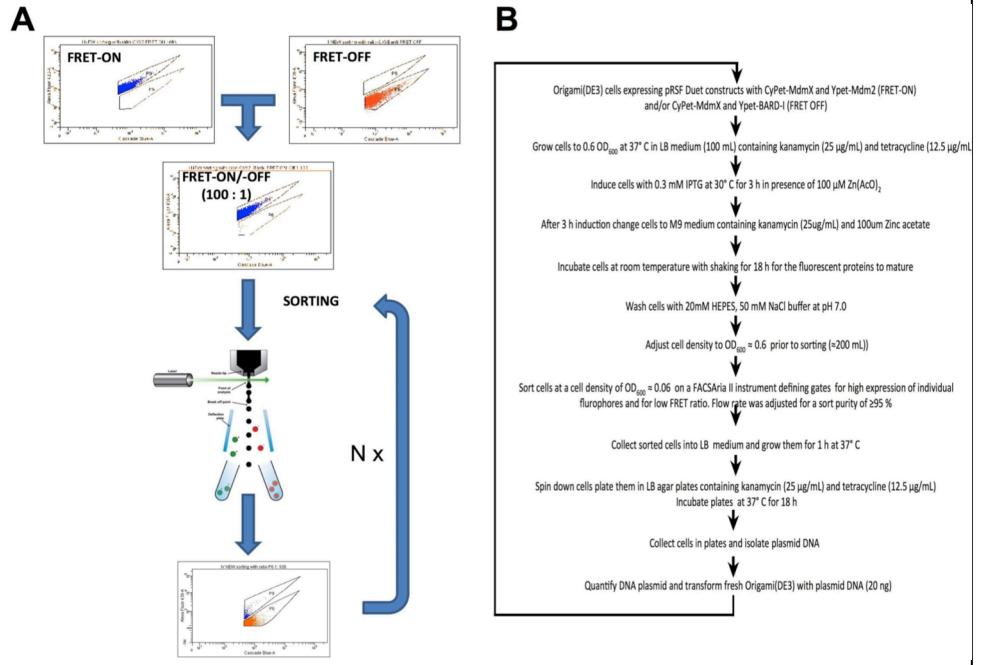


Figure 11. Scheme and protocol used for the sorting of *E. coli* using a FRET-based reporter formed by CyPet-HdmX and YPet-Hdm2.



We are now in the process of testing further this protocol for the enrichment of *E. coli* samples with FRET-OFF/FRET-ON ratios of 1:100,000 and 1:1,000,000. We expect that between 5 and 7 sorting steps will be required to enrich the FRET-OFF phenotype in these mixtures. The validation of this protocol will allow us to start the screening of the first MCoTI-based-libraries containing up to 10^6 different cyclotides within the next month or so. In principle, this approach should be able to enrich genetically-encoded cyclotide-based libraries containing between 1 to 10 hits per 1,000,000 library members.

3) Development of a novel cyclotide with p53 activating properties. Since we do not yet have any rationally engrafted or selected cyclotides and small chemical compounds that could inhibit Hdm2/Hdmx RING interaction, we have used other well-studied protein-protein interactions, i.e., p53/Hdm2 and p53/Hdmx to perform proof-of-principle experiments. During this year we have developed a novel cyclotide able to antagonize the interaction between p53 and Hdm2/HdmX. This cyclotide has been designed and used as model to test the activation of the p53 pathway in vitro and in-cell assays and to test the potential of cyclotide-based therapeutics to target intracellular protein-protein interactions like the RING-mediated Hdm2-HdmX interaction.

During the second year of this project we have developed a cyclotides that activate programmed cellular death, a promising target for cancer therapy [19]. Specifically, we have engineered a cyclotide (MCoTI-PMI) able to inhibit the interaction between p53 and Hdm2/HdmX. This was accomplished by engineering one the loops of the cyclotide MCoTI-I to display a small helical peptide derived from the natural Hdm2-binding sequence of p53 [20]. The α -helical segment was engineered using the peptide apamin [21], a component of *Apis mellifera* venom, to engineer the p53-based α -helical peptide PMI α -helix [20] and a turn into loop 6 of cyclotide MCoTI-I (Fig. 12). Engineered cyclotide MCoTI-PMI was chemically produced (Fig. 13) or recombinantly expressed using intein-mediated cyclization with good yields (Fig. 14). MCoTI-PMI was also shown to fold adopting a native cyclotide fold by 2D-heteronuclear NMR (Fig. 15) and bind with high affinity to both Hdm2 and HdmX with a K_d value of 2.3 ± 0.1 nM and 9 ± 1 nM, respectively as determined by fluorescence anisotropy

Figure 12. Design of MCoTI-PMI, a MCoTI-based cyclotide with the α -helical peptide PMI grafted onto loop 6 of cyclotide MCoTI-I. Binding constants for the different p53-based peptides were taken from reference [1]. The structure of MCoTI-PMI is based on a model built using the structure MCoTI-II and apamin. The grafted sequence is shown in small letters. The sequence of the polypeptide is shown using a single letter code for every amino acid residue.

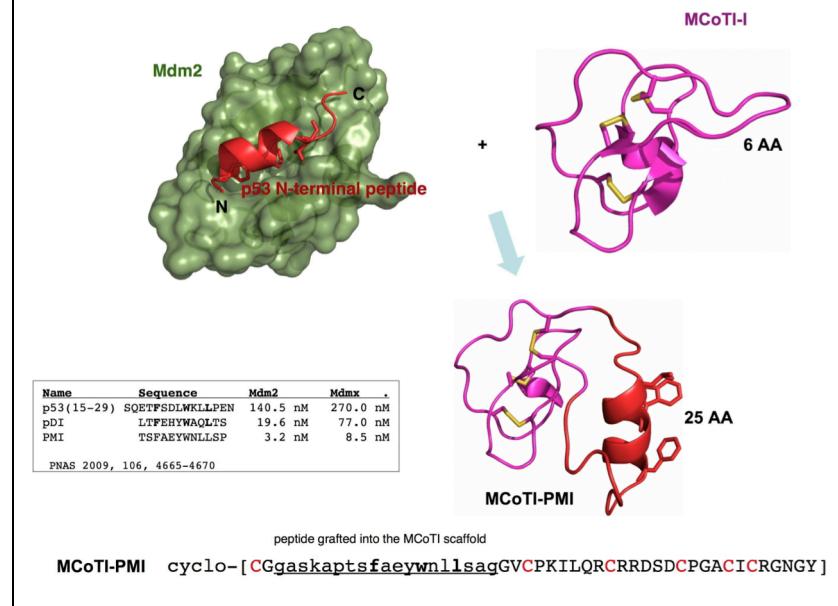
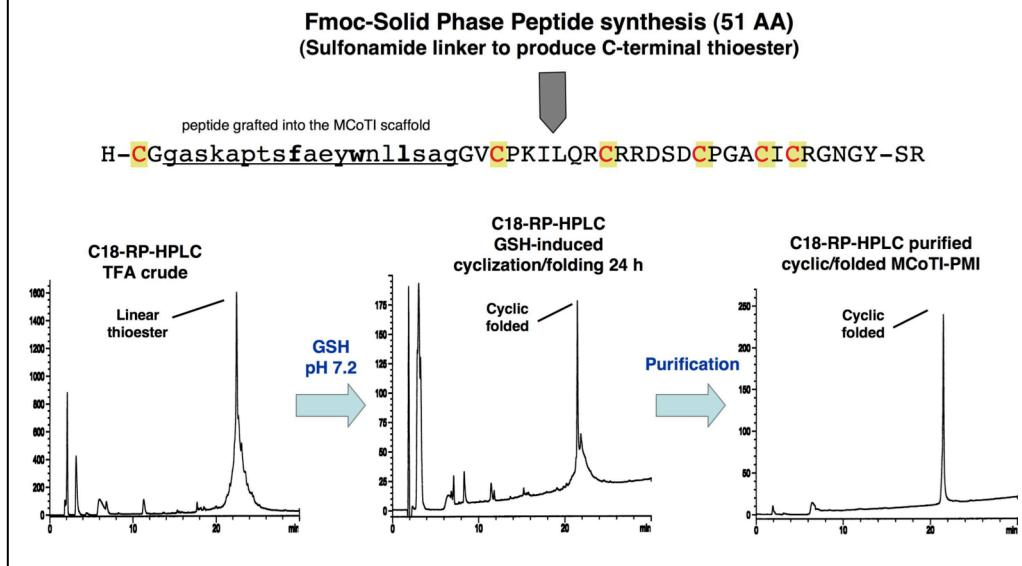


Figure 13. Chemical synthesis of cyclotide MCoTI-PMI using Fmoc-based solid-phase synthesis on a sulfonamide linker. The resulting linear peptide thioester is reacted in PBS in the presence of reduced GSH, this rapidly promotes the backbone cyclization and concomitant folding of the resulting grafted cyclotide MCoTI-PMI.



(Fig. 16). As expected, the cyclotide MCoTI-PMI-F3A mutant did not bind neither the p53 binding domains of Hdm2 nor HdmX. This cyclotide contains a mutation (F3A) in the PMI peptide, which has been shown to disable the interaction between PMI and Hdm2 or HdmX.

Competitive inhibition experiments carried out *in vitro* with a FRET-based reporter formed by the p53 peptide (p53 15-29) and the p53 binding domains of Hdm2 or HdmX fused to the fluorescent protein YPet or CyPet, respectively, also showed that MCoTI-PMI was able to inhibit the p53/Hdm2 and p53/HdmX interaction in a dose-dependent fashion with IC_{50} values of 41 ± 1 nM and 171 ± 1 nM, respectively. The IC_{50} value for nutlin-3 (a small molecule designed to inhibit the p53/Hdm2) [5] was around 3 times larger than of the cyclotide MCoTI-PMI (Fig. 17).

Interestingly, the cyclotide MCoTI-PMI was also able to inhibit the interaction between p53 and HdmX while

Figure 14. Recombinant expression of cyclotide MCoTI-PMI using a protein splicing unit and a TEV protease recognition leading signal. Once the linear precursor is expressed and purified the TEV leading signal is cleaved to yield an N-terminal Cys residue that reacts in an intramolecular fashion with the C-terminal thioester generated by the protein splicing unit or intein. The backbone cyclization is induced by the presence of reduced GSH, which also promotes the folding of the resulting cyclic polypeptide into a cyclotide fold [2].

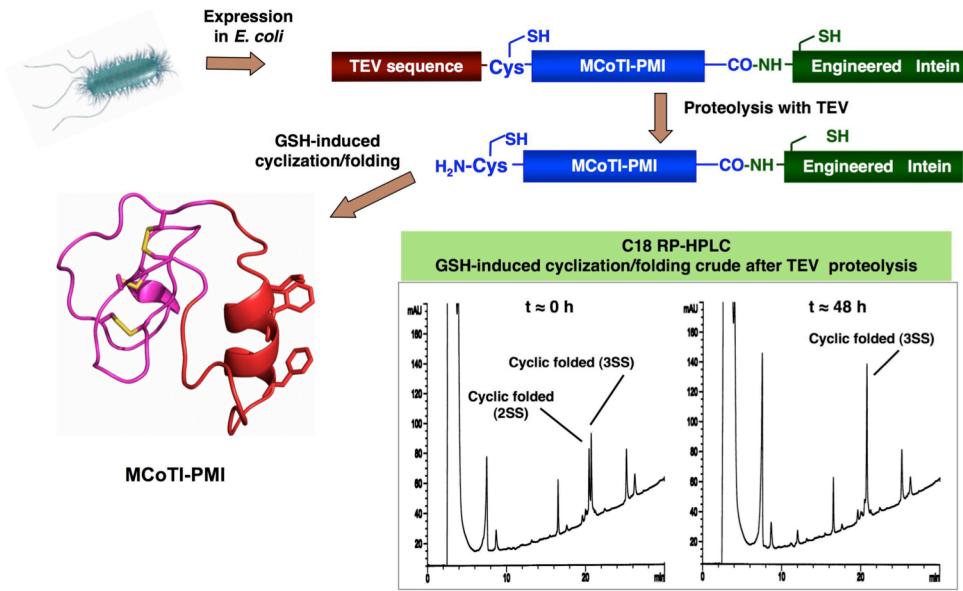
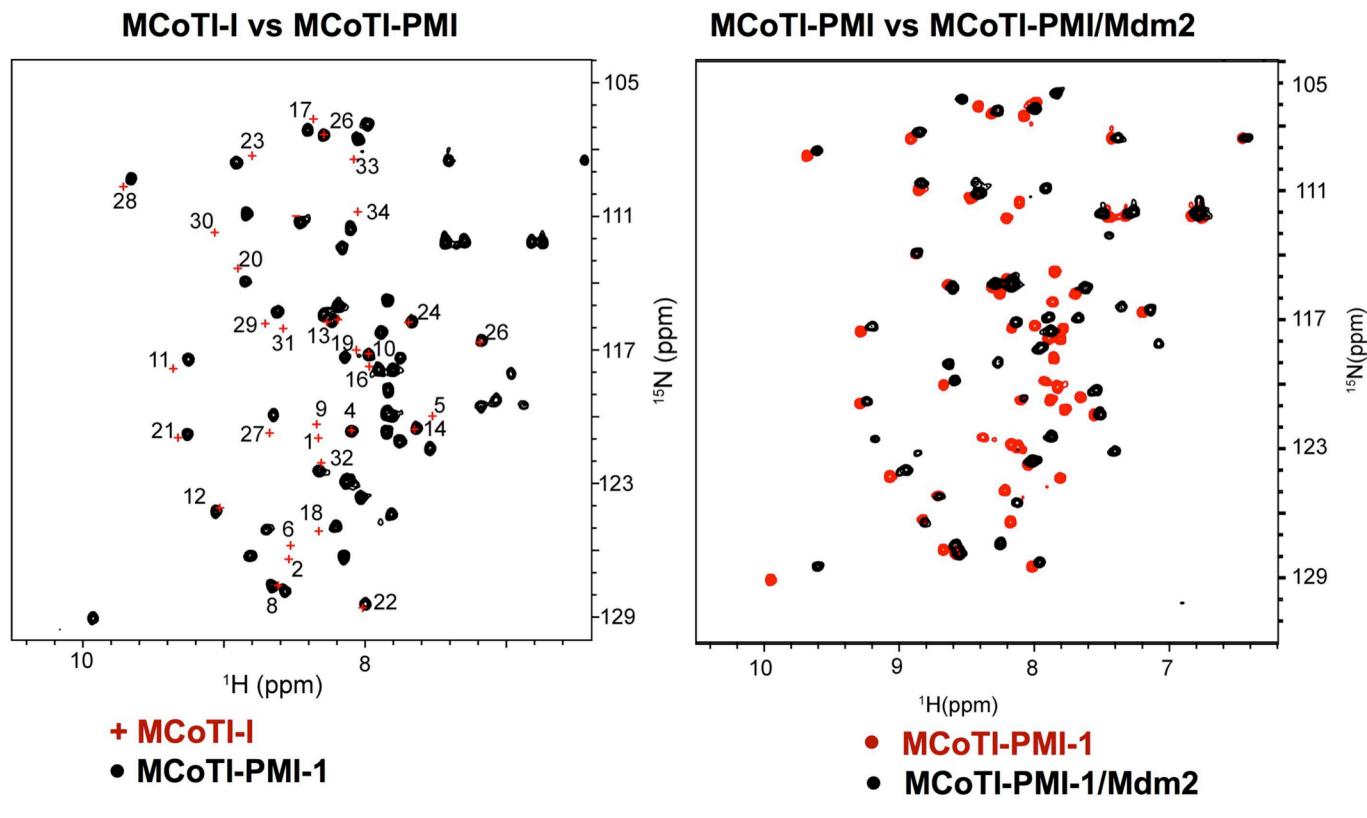
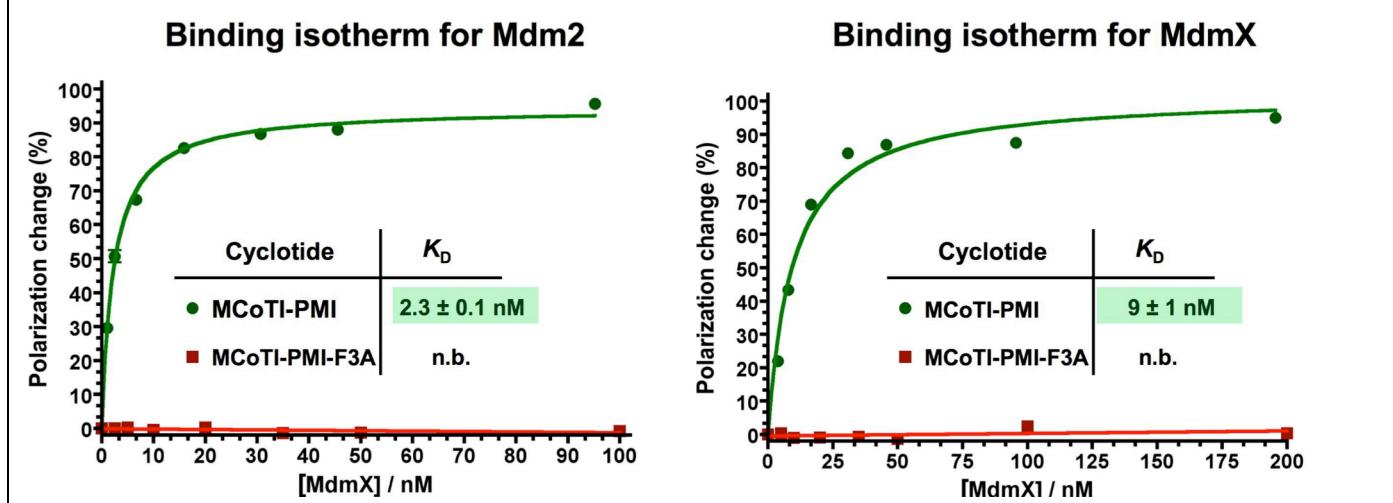


Figure 15. (A) ^{15}N -HSQC spectra of ^{15}N -labeled cyclic MCoTI-I (black) and MCoTI-PMI (red). (B) ^{15}N -HSQC spectra of ^{15}N -labeled MCoTI-PMI free (red) and complexed with the p53 binding domain of Hdm2 (black). All spectra were recorded in sodium phosphate buffer at pH 6.5 at 298 K.



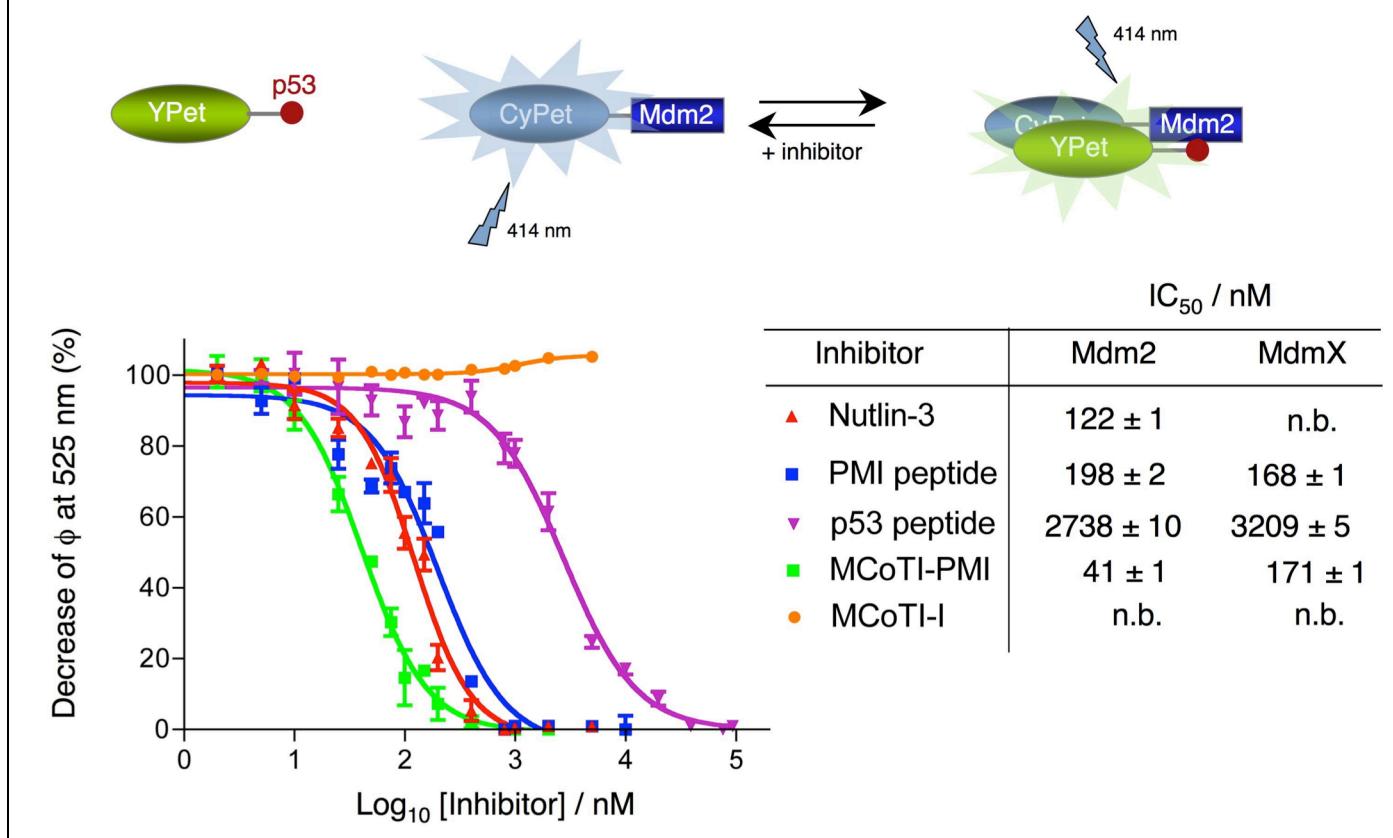
nutlin-3 completely failed to inhibit it. The empty scaffold, i.e. cyclotide MCoTI-I, did not present any activity against either Hdm2 or HdmX (Figs. 16 and 17).

Figure 16. Binding isotherms of the p53 binding domains of Hdm2 and HdmX to cyclotides MCoTI-PMI and MCoTI-PMI-F3A.



More importantly, cyclotide MCoTI-PMI was shown to induce cytotoxicity in p53-wt human cancer cells in a p53 dependent pathway with an IC_{50} value of $\approx 10 \mu M$, which similar to that of the small molecule nutlin-3 recently developed to antagonize p53/Hdm2 [5] (Fig. 18). In contrast to nutlin-3, the cyclotide MCoTI-PMI showed very little toxicity to normal epithelial cell lines (Fig. 18). Importantly, the cyclotide MCoTI-PMI also showed a remarkable resistance to degradation in human serum at $37^\circ C$ ($\tau_{1/2} \approx 32 h$), while the linear and reduced form of MCoTI-PMI was rapidly degraded under the same conditions ($\tau_{1/2} \approx 0.5 h$) (Fig. 19). It is worth

Figure 17. Competitive inhibition of p53/Hdm2 interaction by nutlin-3, different p53-based peptides, cyclotide MCoTI-I and PMI-grafted cyclotide MCoTI-PMI. The experiment was performed using FRET-reporter formed by the N-terminal p53 peptide fused to YPet (YPet-p53) and the p53 binding domain of Hdm2 fused to CyPet (Hdm2-CyPet). Briefly, 20 nM CyPet-Hdm2 and 5 μM YPet-p53 were mixed at $23^\circ C$ in 10 mM sodium phosphate 150 mM NaCl buffer at pH 7.3. Fluorescence was measured at 475 and 525 nm using an excitation wavelength of 414 nm in the presence of different inhibitor concentrations.

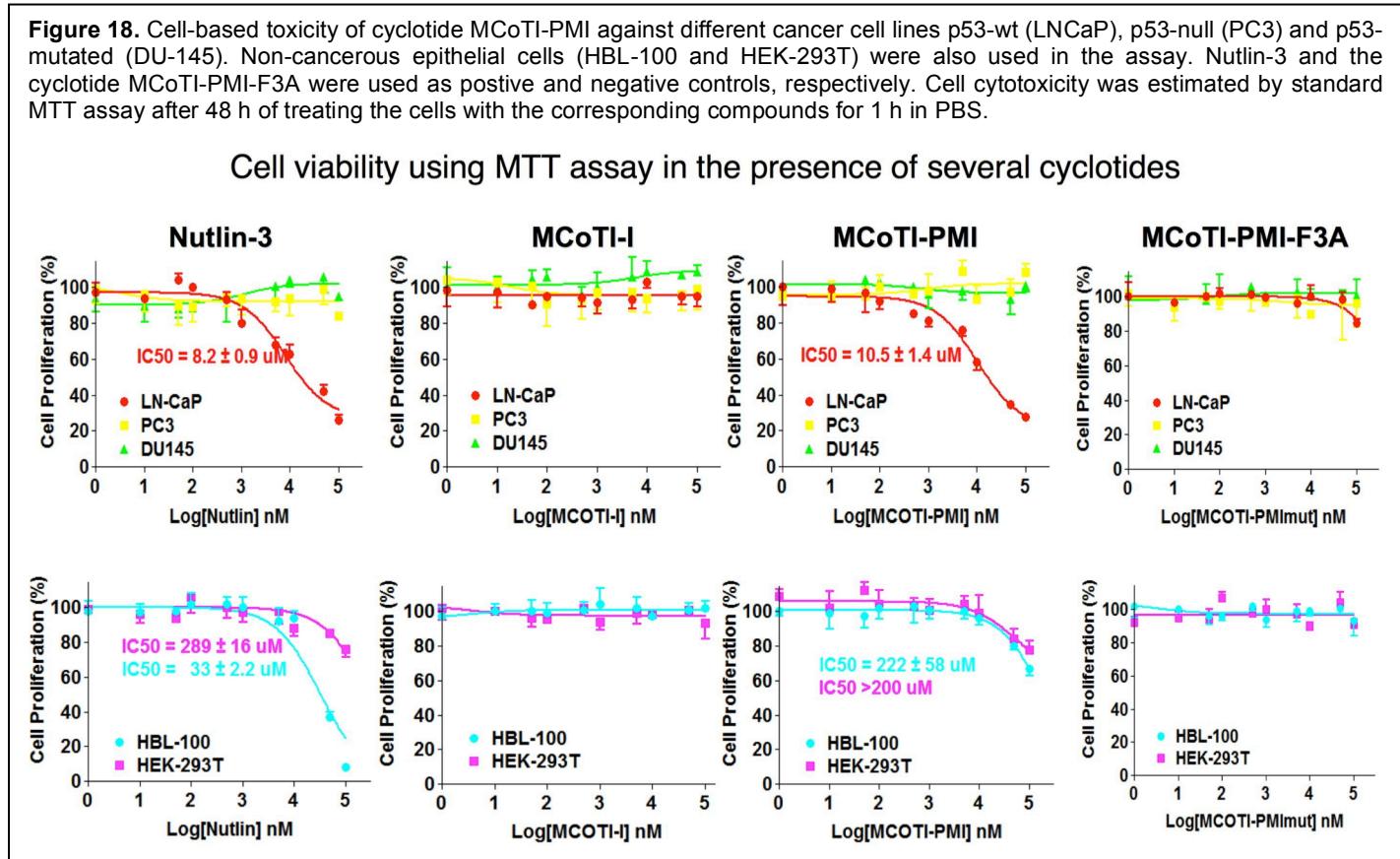


noting that the chemical reduction of MCoTI-PMI required extreme conditions involving the use of 8 M urea in

the presence of 100 mM DTT at pH 8.0. Remarkably, these results demonstrate that MCoTI-based engineered cyclotides can cross cell membranes to target cytosolic protein/protein interactions, such as p53/Hdm2 and p53/HdmX and highlight the extraordinary resistance of this scaffold to chemical, physical and biological degradation.

These studies show the therapeutic potential of cyclotides to target intracellular protein-protein interactions such as the p53-Hdm2/HdmX and the RING-mediated Hdm2-HdmX interaction. It also highlights the abilities of the new USC team to test the biological activity of selected cyclotides to activate the p53 pathway in normal and cancer cells using the MTT assay.

Figure 18. Cell-based toxicity of cyclotide MCoTI-PMI against different cancer cell lines p53-wt (LNCaP), p53-null (PC3) and p53-mutated (DU-145). Non-cancerous epithelial cells (HBL-100 and HEK-293T) were also used in the assay. Nutlin-3 and the cyclotide MCoTI-PMI-F3A were used as positive and negative controls, respectively. Cell cytotoxicity was estimated by standard MTT assay after 48 h of treating the cells with the corresponding compounds for 1 h in PBS.

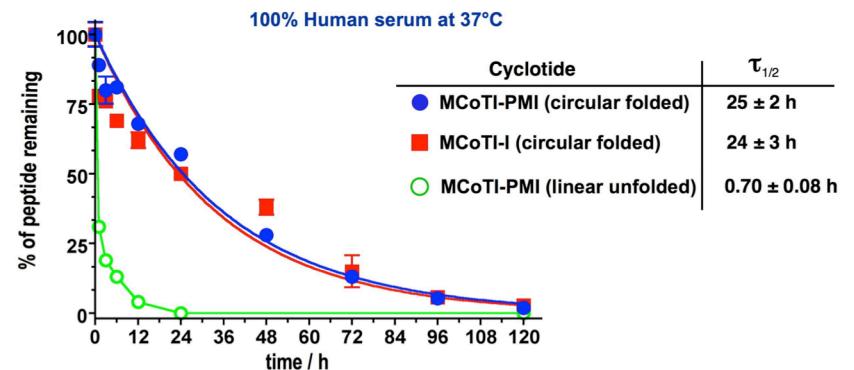


4) Development of mammalian cell-based assay for testing biological activity of cyclotide-based Hdm2/Hdmx interaction antagonists.

This work was performed by Dr. Wahl as a subcontract to this project. The development of mammalian cell-based Hdm2/Hdmx protein-protein interaction assay has been completed. We utilized *bimolecular luciferase complementation* (BiLC) to achieve this goal. Originally developed by Luker et al., BiLC relies on the reconstitution of luciferase activity from two fragments of the luciferase cDNA [22]. These halves are fused to proteins of interest resulting in luciferase activity only if the two test proteins interact. We constructed fusion proteins of the Hdmx RING (amino acids 427-490) and Hdm2 RING (amino acids 429-491) domains with N- (amino acids 1-416) and C- (amino acids 398-550) terminal fragments of firefly luciferase (designated nLuc-XR and cLuc-2R, respectively). A RING domain mutant of Hdm2 (C464A) fused to the cLuc

Figure 19. Serum stability of cyclotide MCoTI-PMI in human serum at 37° C. Cyclotide MCoTI-I and linearized/reduced MCoTI-PMI were also used as controls.

Serum stability of MCoTI-I and grafted MCoTI-cyclotides



fragment (cLuc-2R_C464A) serves as a negative control, as this mutant is unable to heterodimerize with Hdmx. The constructs are expressed from a bidirectional doxycycline/tetracycline-inducible promoter (TRE_{bi}), affording precise control over the timing and expression level of $\frac{1}{2}$ Luc-RING fusions. The DNA constructs are depicted in Figure 20A. Using recombinase-mediated cassette exchange [23], we introduced this construct into p53-null SaOS-2 osteosarcoma cells, and into p53 wild type U2OS osteosarcoma cells at a single site in the genome of each cell line. The cell clones have been isolated that show highly reproducible induction with 10-800 fold dynamic range for generating luminescence (Figure 20B). Importantly, luciferase activity is specifically

Figure 20. A. Diagram of Hdm2/Hdmx RING domain BiLC expression cassette. **B.** The Hdm2/Hdm4 RING domain BiLC signals are doxycycline dose dependent. Saos-2 cells expressing the nLuc-XR/cLuc-2R showed robust and doxycycline dose dependent luminescence in an end-point luciferase assay (green bars) whereas Hdm2 C464A RING mutant C464A generated no luminescent signal (magenta bars).

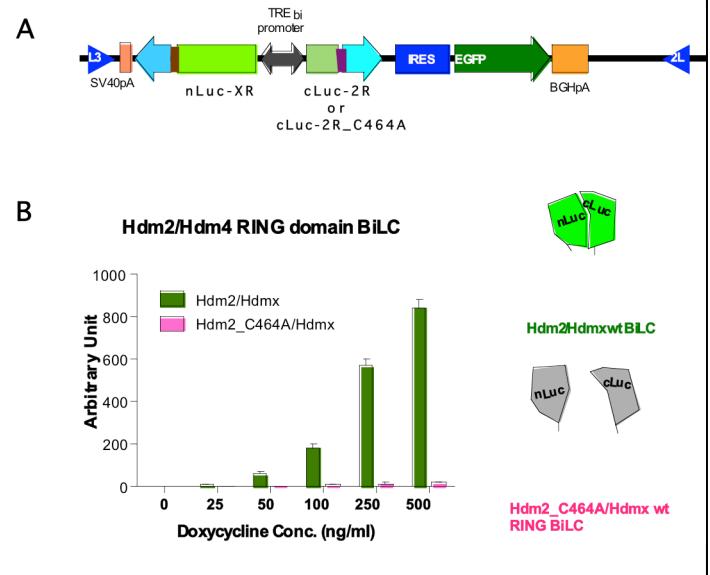
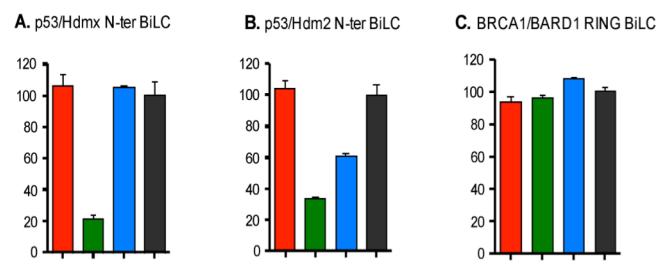


Figure 21. In vitro and in vivo BiLC assays. Cellular lysates prepared from U2OS p53/Hdmx (**A**), p53/Hdm2 (**B**) and BRCA1/BARD1 RING domain (**C**) BiLC reporter cell lines were incubated with MCoTI-I (10 μ M), MCoTI-PMI (10 μ M), Nutlin-3a (10 μ M), and DMSO in the presence of D-luciferin reaction buffer [2] in a 384-well plate (Corning solid white). Luminescence was read in a Tecan M200 plate reader with 0.5 sec integration time per well at 26 °C. The percentage of luminescence is expressed as the luminescent reading from each treatment over that of DMSO. Each treatment contained three duplicates and error bar indicated SEM.



dependent on the interaction of the Hdm2 and Hdmx RING domains, since only background luminescence was observed upon induction of the gene encoding the $\frac{1}{2}$ luciferase Hdm2_C464A mutation (Fig. 20B). We had completed these BiLC reporter cell lines. We anticipate that any cyclotide antagonists that can inhibit Hdm2/Hdmx RING interaction will decrease luminescence signal in BiLC assay. Although currently we do not have any specific positive compound to inhibit Hdm2/Hdmx RING interaction, we have demonstrated that this BiLC strategy is able to identify protein-protein interaction antagonists in other protein-protein interaction BiLC system using the interaction between p53 and Hdm2/HdmX and the cyclotide MCoTI-PMI (Fig. 21).

5) Biosynthesis of other circular polypeptides in *E. coli*. During this second year of the project we have also explored the biosynthesis of other circular polypeptides (circular α -defensins and naturally occurring circular θ -defensins) using protein splicing in *E. coli* cells. The purpose was to explore the generality of the circularization method with other scaffolds that could be also used potentially for the generation of libraries to screen the Hdm2-HdmX interaction. The data obtained in this section is summarized in papers #3, #4 and #5 (see appendix section).

Key Research Accomplishments

- We have accomplished the biosynthesis of a large combinatorial library ($\approx 10^6$ members) of genetically-encoded cyclotides in *Escherichia coli* using loop 2 of cyclotide MCoTI-I as molecular template. This is the first time that a library with such a high diversity has been produced. This library has been used to screen RING Mdm2/MdmX antagonists.
- We have accomplished of cyclotide-based libraries into different *E. coli* expression plasmids with different promoters, origin of replication and antibiotic resistance to allow the screening of the libraries using cell-

based reporters encoded in orthogonal plasmids. This is required to allow the in-cell screening of genetically encoded cyclotide-based libraries

- We have carried out the first ever reported study of the backbone dynamics of a natively folded MCoTI-I cyclotide in the free state and complexed to its binding partner trypsin in solution. This accomplishment is critical to understand the dynamics and better define the loops that are best amenable for randomization in the generation of MCoTI-based libraries (Appendix: paper #8)
- We have optimized the biosynthesis of cyclotides in living cells using PTS-mediated backbone cyclization. Using this new methodology we can reach intracellular concentrations of folded cyclotides up to 7 μ M. This approach can be also used for the incorporation of unnatural amino acids (UAs), which opens the exciting possibility for the generation of genetically-encoded cyclotide-based libraries containing additional chemical diversity for selection on novel cyclotide-based sequences with improved biological activity (Appendix: paper #7)
- We have developed a protocol for phenotypic sorting of *E. coli* cells using FACS in combination with the FRET-based reporter to screen for RING-mediated Hdm2-HdmX interactions. We have tested the protocol to enrich samples from 1:100 (FRET-OFF:FRET-ON) to 100:1 (FRET-OFF:FRET-ON). This can be done in just 4 sorting steps.
- Since we do not yet have any rationally engrafted or selected cyclotides and small chemical compounds that could inhibit Hdm2-HdmX RING interaction, we used other well-studied protein-protein interactions, i.e., p53-Hdm2 and p53-HdmX to perform proof-of-principle experiments. Therefore, we have developed a cyclotide (MCoTI-PMI) able to antagonize the p53-Hdm2 and p53-HdmX. This novel cyclotide was biologically active in promoting cytotoxicity in p53-wt prostate cancer cell lines (LnCaP) and showed low toxicity to normal epithelial cells. The cyclotide MCoTI-PMI was fully characterized by NMR showing a well-folded structure and a remarkable stability to biological degradation in human serum (Manuscript in preparation).
- Developed mammalian and in-vitro and in-cell based Hdm2/Hmdx RING domain interaction BiLC assay. This work was performed by the Wahl group under a subcontract to this project.

Reportable outcomes

Peer-reviewed Publications Submitted and Published:

- J. J. Contreras, A. Y. O. Elnagar, S. Hamm-Alvarez and **J. A. Camarero** (2011) Cellular Uptake of cyclotide MCoTI-I follows multiple endocytic pathways, *J. Control Release*, 115(2), 134-143 (Appendix: paper #1).
- A. Gould, Y. Li, T. L. Aboye and **J. A. Camarero** (2011) Cyclotides, a novel ultrastable polypeptide scaffold for drug discovery, *Curr. Pharm. Des.*, 17(38), 4294-4307 (Appendix: paper #2).
- A. E. Garcia, K. P. Tai, S. S. Puttamadappa, A. Shekhtman, A. J. Ouellette and **J. A. Camarero** (2011) Biosynthesis and antimicrobial evaluation of backbone-cyclized alpha-defensins, *Biochemistry*, 50(48), 10508-10519 (Appendix: paper #3).
- A. Gould, Y. Li, S. Majumder, A. E. Garcia, P. Carlsson, A. Shekhtman and **J. A. Camarero** (2012) Recombinant production of rhesus θ -defensin-1 (RTD-1) using a bacterial expression system, *Mol. Biosyst.*, 8(4), 1359-1365 (Appendix: paper #4).
- T. L. Aboye, Y. Li, S. Majumder, A. Shekhtman and **J. A. Camarero** (2012) Efficient one-pot cyclization/folding of Rhesus θ -defensin-1 (RTD-1), *Bioorg. Med. Chem. Lett.*, 22(8):2823-2826 (Appendix: paper #5).
- T. L. Aboye and **J. A. Camarero** (2012) Biological synthesis of circular proteins, *J. Biol. Chem.*, 287(32), 27026-27032. (Appendix: paper #6).
- K. Jagadish, R. Borra, V. Lacey, S. Majumder, A. Shekhtman, L. Wang and **J. A. Camarero** (2013) Expression of fluorescent cyclotides using protein trans-splicing for easy monitoring of cyclotide-protein interactions, *Angew. Chem. Int. Ed.*, 52(11), 3037-3279 (Appendix: paper #7).

- S. S. Puttamadappa, K. Jagadish, A. Shekhtman and **J. A. Camarero** (2010) Backbone dynamics of cyclotide MCoTI-I free and complexed with trypsin, *Angew. Chem. Int. Ed.*, **49**(39), 7030-7034 (Appendix: paper #8).
- A. E. Garcia and **J. A. Camarero** (2010) Biological Activities of Natural and Engineered Cyclotides, a Novel Molecular Scaffold for Peptide-Based Therapeutics, *Curr. Mol. Pharmacol.*, **3**(3), 153-163 (Appendix: paper #9).
- **J. A. Camarero** (2011) Legume cyclotides shed new light on the genetic origin of knotted circular proteins, *Proc. Natl. Acad. Sci. USA*, **108**(25), 10025-10026 (Appendix: paper #10)

Oral presentations:

- Constrained Peptides and Macrocylics Drug Discovery – Novel Peptide Therapeutics: Targeting protein-protein interactions with engineered cyclotides, April 16-17, 2012, San Diego, California.
- PepTalk 2013. Choosing, Designing and Optimizing Hosts and Platforms. Methods, Strategies and Applications: Efficient Biosynthesis of Native Cyclotides Using Intein-Mediated Protein Trans-Splicing, January 24-25, 2013, Palm Springs, California.
- 13th Tetrahedron Symposium – Asia Edition. Challenges in Bioorganic and Organic Medicinal Chemistry: Efficient biosynthesis of cyclotides inside living cells using intein-mediated protein trans-splicing, November 27-30, 2012, Taipei, Taiwan.
- Invited talk to the seminar series of the Department of Pharmacy, National Taiwan University: Cyclotides, a novel scaffold for targeting protein-protein interactions, November 29, 2012, Taipei, Taiwan.
- Invited talk to the 2nd International Conference on Circular Proteins: Targeting protein-protein interactions with engineered cyclotides, October 14-17, 2012, Heron Island, Gladstone, Queensland, Australia.
- Invited talk to AgraQuest-Bayer: Cyclotides: a novel source of agrochemicals, September 21, 2012, Davis, New Jersey.
- Invited talk to the Gordon Research Conference on Medicinal Chemistry, New London, New Hampshire, August 5-10, 2012.
- Discovery and Development of Novel Therapeutics: Recent Advances and Future Trends, Norris Cancer Center, University of Southern California, Los Angeles, May 18, 2012, Los Angeles, California.
- PepTalk 2012. Recombinant Protein Therapeutics: Targeting Protein/Protein Interactions, January 9-11, 2012, San Diego, California.
- Invited talk to 2011 ACS 43rd Western Regional Meeting Program: Biomolecular engineering of drug carriers session, November 10-12, 2011, Pasadena, California.
- Invited talk to Bristol-Myers Squibb: Drug Discovery Seminar Series, October 25-26, 2011, Princeton, New Jersey.
- Seventh Annual PEGS (Protein Engineering Summit) Conference, Phage and Yeast Display of Antibodies and Proteins Session, May 9-13, 2011, Boston.
- 2011 Spring ACS National Meeting, Division of Biological Chemistry, Invited Lecture to the Ralph F. Hirschmann Award in Peptide Chemistry: Symposium in Honor of David J. Craik, March 28, 2011, Anaheim, California.
- University of Uppsala, Invited seminar to The Svedberg Lecture Series, March 24, 2011, Uppsala, Sweden.
- Pacificchem 2010: International Chemical Congress of Pacific Basic Societies, December 15-20, Honolulu, Hawaii.
- Roche Colorado Corporation Peptide Symposium (RCCPS) 2010, Cyclotides, a novel natural peptide scaffold for drug discovery, September 14-16, Boulder, Colorado.
- Natural Peptides to Drugs (NP2D) 4th International Congress, April 11-14, 2010, Zermatt, Switzerland.
- BIT's 3rd Annual Protein and Peptide Conference (PepCon), March 21-23, 2010, Beijing, China.

- First Look LA, Cyclopeptics: Providing a new class of therapeutics for the treatment of human disease, November 4, 2009, California Institute of Technology, Pasadena, USA.
- Barcelona BioMed Conference Peptide Engineering: Therapeutic Peptides, 26-28 October, 2009, Barcelona, Spain.
- 1st International conference on circular proteins at the 8th Australian Peptide Conference, 18-21 October, 2009, Heron Island, Australia.
- 3rd Annual PEGS (Protein Engineering Summit) Europe Conference, 1st Next generation technologies for protein science, October 6-8, 2009, Hannover, Germany.
- ACS 238th National Meeting, Biotechnological Solutions to Global Security: Antibiotics for Combating Infectious Diseases Session, August 16-20, 2009, Washington.
- 11th International Congress on Amino Acids, Peptides and Proteins. Peptides-Chemistry and Biology Session, August 3-7, 2009, Vienna, Austria.
- Fourth Annual PEGS (Protein Engineering Summit) Conference, 4th Proteins Difficult to Express Session, April 28-May 2, 2008, Boston.

Reagents developed in the Wahl group:

- Four reporter mammalian cell lines:
 1. Saos-2 134-283: p53-null; Hdm2/Hdmx RING domain BiLC reporter cell line.
 2. Saos-2 134-285: p53-null; Hdm2_C464A/Hmdx RING domain BiLC reporter cell line.
 3. U2OS 134-283: p53-wt; Hdm2/Hdmx RING domain BiLC reporter cell line.
 4. U2OS 134-285: p53-wt; Hdm2_C464A/Hmdx RING domain BiLC reporter cell line.

Conclusion

The results accomplished during the 2 year of the proposal are extremely encouraging. We have shown that protein trans-splicing can be used for efficient production cyclotides in living cells reaching intracellular concentrations up to 7 μ M, and more importantly allows the efficient incorporation of unnatural amino acids (UAs) to improve the chemical diversity of the biologically-generated libraries of cyclotides as well as their pharmacological properties. For example the inclusion of amino acids with side-chains containing classical warheads as ketoamides, boronates and hydroxamates, which are absent from the 20 proteinogenic amino acids, should help in screening and selection of cyclotides with more effective pharmacological properties.

We have developed FRET-based fluorescence reporter that can be used for monitor inhibition of the RING-mediated Hdm2/HdmX interaction inside living cells using high throughput cell-sorting techniques. During the second year we have developed a FACS-based protocol to screen and sort cells based on their FRET-phenotype. For example we have enriched in just 4 sorting steps a library of cells (99:1 FRET-ON:FRET-OFF) into cells having the FRET-OFF phenotype. During the last year of the project we will use this protocol to enrich cyclotide-based libraries for its ability to antagonize the RING-mediated Hdm2-HdmX interaction.

Since we do not yet have any rationally engrafted or selected cyclotides and small chemical compounds that could inhibit Hdm2-Hdmx RING interaction, we have used other well-studied protein-protein interactions, i.e., p53-Hdm2 and p53-Hdmx to perform proof-of-principle experiments. Hence, we have developed a novel cyclotide able to antagonize the p53-Hdm2 and p53-HdmX interaction in vitro and in-cell based experiments using a panel of different prostate cancer cells including p53-wt (LnCap), p53-mutated (DU145) and p53-null (PC3) as well a normal epithelial cells. Importantly, this cyclotide showed activity similar to the small molecule p53-Hdm2 antagonist nutlin-3. These experiments were carried out in the Camarero lab and will be also used to evaluate any cyclotide selected by molecular evolution able to antagonize the Hdm2-HdmX interaction and therefore activate the p53 pathway.

The Wahl group has confirmed these results by using an in-house developed BiLC system. Unfortunately, Dr.

Wahl has discontinued his collaboration due to a conflict of interest. During the 3rd year of the project the biological activity of selected cyclotides will be tested in the Camarero Lab as described earlier for the cyclotide MCo-PMI.

Altogether these results validate the technologies required for the cell-based selection of cyclotides able to antagonize the Hdm2-HdmX interaction, and it will make possible to carry out the selection of new cyclotides able to antagonize this RING-mediated interaction in the near future.

References

- [1] Pazgier, M.;Liu, M.;Zou, G.;Yuan, W.;Li, C.;Li, J.;Monbo, J.;Zella, D.;Tarasov, S.G.;Lu, W. Structural basis for high-affinity peptide inhibition of p53 interactions with MDM2 and MDMX. *Proc Natl Acad Sci U S A*, **2009**, *106*, 4665-4670.
- [2] Sancheti, H.;Camarero, J.A. "Splicing up" drug discovery. Cell-based expression and screening of genetically-encoded libraries of backbone-cyclized polypeptides. *Adv Drug Deliv Rev*, **2009**, *61*, 908-917.
- [3] Zhang, Z.;Li, M.;Wang, H.;Agrawal, S.;Zhang, R. Antisense therapy targeting MDM2 oncogene in prostate cancer: Effects on proliferation, apoptosis, multiple gene expression, and chemotherapy. *Proc Natl Acad Sci U S A*, **2003**, *100*, 11636-11641.
- [4] Logan, I.R.;McNeill, H.V.;Cook, S.;Lu, X.;Lunec, J.;Robson, C.N. Analysis of the MDM2 antagonist nutlin-3 in human prostate cancer cells. *Prostate*, **2007**, *67*, 900-906.
- [5] Vassilev, L.T.;Vu, B.T.;Graves, B.;Carvajal, D.;Podlaski, F.;Filipovic, Z.;Kong, N.;Kammlott, U.;Lukacs, C.;Klein, C.;Fotouhi, N.;Liu, E.A. In vivo activation of the p53 pathway by small-molecule antagonists of MDM2. *Science*, **2004**, *303*, 844-848.
- [6] Vousden, K.H.;Lane, D.P. p53 in health and disease. *Nat Rev Mol Cell Biol*, **2007**, *8*, 275-283.
- [7] Kimura, R.H.;Steenblock, E.R.;Camarero, J.A. Development of a cell-based fluorescence resonance energy transfer reporter for *Bacillus anthracis* lethal factor protease. *Anal Biochem*, **2007**, *369*, 60-70.
- [8] Camarero, J.A.;Kimura, R.H.;Woo, Y.H.;Shekhtman, A.;Cantor, J. Biosynthesis of a fully functional cyclotide inside living bacterial cells. *Chembiochem*, **2007**, *8*, 1363-1366.
- [9] Austin, J.;Kimura, R.H.;Woo, Y.H.;Camarero, J.A. In vivo biosynthesis of an Ala-scan library based on the cyclic peptide SFTI-1. *Amino Acids*, **2010**, *38*, 1313-1322.
- [10] Austin, J.;Wang, W.;Puttamadappa, S.;Shekhtman, A.;Camarero, J.A. Biosynthesis and biological screening of a genetically encoded library based on the cyclotide MCoT-I. *Chembiochem*, **2009**, *10*, 2663-2670.
- [11] Young, T.S.;Young, D.D.;Ahmad, I.;Louis, J.M.;Benkovic, S.J.;Schultz, P.G. Evolution of cyclic peptide protease inhibitors. *Proc Natl Acad Sci U S A*, **2011**, *108*, 11052-11056.
- [12] Scott, C.P.;Abel-Santos, E.;Wall, M.;Wahnon, D.;Benkovic, S.J. Production of cyclic peptides and proteins in vivo. *Proc. Natl. Acad. Sci. USA*, **1999**, *96*, 13638-13643.
- [13] Tavassoli, A.;Benkovic, S.J. Split-intein mediated circular ligation used in the synthesis of cyclic peptide libraries in *E. coli*. *Nat Protoc*, **2007**, *2*, 1126-1133.
- [14] Zettler, J.;Schutz, V.;Mootz, H.D. The naturally split Npu DnaE intein exhibits an extraordinarily high rate in the protein trans-splicing reaction. *FEBS Lett*, **2009**, *583*, 909-914.
- [15] Iwai, H.;Zuger, S.;Jin, J.;Tam, P.H. Highly efficient protein trans-splicing by a naturally split DnaE intein from *Nostoc punctiforme*. *FEBS Lett*, **2006**, *580*, 1853-1858.
- [16] Shah, N.H.;Vila-Perello, M.;Muir, T.W. Kinetic control of one-pot trans-splicing reactions by using a wild-type and designed split intein. *Angew Chem Int Ed Engl*, **2011**, *50*, 6511-6515.
- [17] Kwon, Y.;Coleman, M.A.;Camarero, J.A. Selective Immobilization of Proteins onto Solid Supports through Split-Intein-Mediated Protein Trans-Splicing. *Angew. Chem. Int. Ed.*, **2006**, *45*, 1726-1729.
- [18] Wang, L.;Xie, J.;Schultz, P.G. Expanding the genetic code. *Annu Rev Biophys Biomol Struct*, **2006**, *35*, 225-249.
- [19] Wang, X. p53 regulation: Teamwork between RING domains of Mdm2 and MdmX. *Cell Cycle*, **2011**, *10*, 4225-4229.
- [20] Li, C.;Pazgier, M.;Yuan, W.;Liu, M.;Wei, G.;Lu, W.Y.;Lu, W. Systematic mutational analysis of peptide inhibition of the p53-MDM2/MDMX interactions. *J Mol Biol*, **2010**, *398*, 200-213.
- [21] Li, C.;Pazgier, M.;Liu, M.;Lu, W.Y.;Lu, W. Apamin as a template for structure-based rational design of potent peptide activators of p53. *Angew Chem Int Ed Engl*, **2009**, *48*, 8712-8715.
- [22] Luker, K.E.;Smith, M.C.;Luker, G.D.;Gammon, S.T.;Piwnica-Worms, H.;Piwnica-Worms, D. Kinetics of regulated protein-protein interactions revealed with firefly luciferase complementation imaging in cells and living animals. *Proc Natl Acad Sci U S A*, **2004**, *101*, 12288-12293.
- [23] Wong, E.T.;Kolman, J.L.;Li, Y.C.;Mesner, L.D.;Hillen, W.;Berens, C.;Wahl, G.M. Reproducible doxycycline-inducible transgene expression at specific loci generated by Cre-recombinase mediated cassette exchange. *Nucleic Acids Res*, **2005**, *33*, e147.

Abbreviations

BiLC, bimolecular luciferase complementation; ES-MS, electrospray mass spectrometry; EPL, expressed protein ligation; FACS, fluorescence-activated cell sorting; FRET, Foster resonance emission transfer; HSQC, heteronuclear single quantum coherence; GSH, reduced glutathione; NMR, nuclear magnetic resonance; PTS, protein trans-splicing; SPPS, solid-phase peptide synthesis; TEV, tobacco etch virus; UAA, unnatural amino acid;

Appendix



Cellular uptake of cyclotide MCoTI-I follows multiple endocytic pathways

Janette Contreras ¹, Ahmed Y.O. Elnagar ¹, Sarah F. Hamm-Alvarez, Julio A. Camarero *

Department of Pharmacology and Pharmaceutical Sciences, School of Pharmacy, University of Southern California, Los Angeles, CA 90033, USA

ARTICLE INFO

Article history:

Received 10 May 2011

Accepted 23 August 2011

Available online 30 August 2011

Keywords:

Cyclotides

Cell penetrating peptides

Endocytosis

Peptide scaffold

ABSTRACT

Cyclotides are plant-derived proteins that naturally exhibit various biological activities and whose unique cyclic structure makes them remarkably stable and resistant to denaturation or degradation. These attributes, among others, make them ideally suited for use as drug development tools. This study investigated the cellular uptake of cyclotide, MCoTI-I in live HeLa cells. Using real time confocal fluorescence microscopy imaging, we show that MCoTI-I is readily internalized in live HeLa cells and that its endocytosis is temperature-dependent. Endocytosis of MCoTI-I in HeLa cells is achieved primarily through fluid-phase endocytosis, as evidenced by its significant colocalization with 10K-dextran, but also through other pathways as well, as evidenced by its colocalization with markers for cholesterol-dependent and clathrin-mediated endocytosis, cholera toxin B and EGF respectively. Uptake does not appear to occur only via macropinocytosis as inhibition of this pathway by Latrunculin B-induced disassembly of actin filaments did not affect MCoTI-I uptake and treatment with EIPA which also seemed to inhibit other pathways collectively inhibited approximately 80% of cellular uptake. As well, a significant amount of MCoTI-I accumulates in late endosomal and lysosomal compartments and MCoTI-I-containing vesicles continue to exhibit directed movements. These findings demonstrate internalization of MCoTI-I through multiple endocytic pathways that are dominant in the cell type investigated, suggesting that this cyclotide has ready access to general endosomal/lysosomal pathways but could readily be re-targeted to specific receptors through addition of targeting ligands.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

Cyclotides are fascinating micro-proteins ranging from 28 to 37 amino acid residues that are naturally expressed in plants and exhibit various biological activities such as anti-microbial, insecticidal, cytotoxic, antiviral (against HIV), and protease inhibitory activity, as well as exert hormone-like effects [1–4]. They share a unique head-to-tail circular knotted topology of three disulfide bridges, with one disulfide penetrating through a macrocycle formed by the other two disulfides and inter-connecting peptide backbones, forming what is called a cystine knot topology (Fig. 1). This cyclic cystine knot (CCK)

framework gives the cyclotides exceptional resistance to thermal and chemical denaturation, and enzymatic degradation [4,5]. In fact, the use of cyclotide-containing plants in indigenous medicine first highlighted the fact that the peptides are resistant to boiling and are apparently orally bioavailable [6].

Cyclotides have been isolated from plants in the *Rubiaceae*, *Violaceae*, *Cucurbitaceae* [4,7] and most recently *Fabaceae* families [8]. Around 160 different cyclotide sequences have been reported in the literature [9,10], although it has been estimated that ≈50,000 cyclotides might exist [11,12]. Despite the sequence diversity, all cyclotides share the same CCK motif (Fig. 1). Hence, these micro-proteins can be considered natural combinatorial peptide libraries structurally constrained by the cystine-knot scaffold [2] and head-to-tail cyclization, but in which hypermutation of essentially all residues is permitted with the exception of the strictly conserved cysteines that comprise the knot.

Cyclotides are ribosomally produced in plants from precursors that comprise between one and three cyclotide domains. However, the mechanism of excision of the cyclotide domains and ligation of the free N- and C-termini to produce the circular peptides has not been completely elucidated yet. It is suspected, however, that specific asparaginyl endopeptidases are involved in the proteolytic processing and cyclization of the precursor proteins [13–15]. Cyclotides can also be produced chemically using solid-phase peptide synthesis in combination with native chemical ligation [16–19], or recombinantly in bacteria by using modified protein splicing units or inteins [20,21].

Abbreviations: Boc, tert-butyloxycarbonyl; CCK, cyclic cystine knot; CPPs, cell-penetrating peptides; CTX-B, cholera toxin B; DAST, diethylaminosulfur trifluoride; DCM, dichloromethane; 10K-Dex, 10,000 MW-dextran; DIEA, N,N-disopropylethylamine; DMF, dimethyl formamide; EGF, epidermal growth factor; EIPA, 5-(N-ethyl-N-isopropyl)amiloride; ES-MS, electrospray-mass spectrometry; Fmoc, 9-fluorenyloxy carbonyl; HBTU, 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate; HPLC, high performance liquid chromatography; Lat B, latrunculin B; MCoTI, *Momordica cochinchinensis* trypsin inhibitor; NMP, N-methyl-pyrrolidone; NMR, nuclear magnetic resonance; RFP-Lamp1, Red Fluorescent Protein-lysosomal associated protein 1; RP-HPLC, reverse phase-high performance liquid chromatography; TFA, trifluoroacetic acid; TIS, tri-isopropylsilane; TR, Texas Red; Trt, trityl.

* Corresponding author at: Department of Pharmacology and Pharmaceutical Sciences, University of Southern California, 1985 Zonal Avenue, PSC 616, Los Angeles, CA 90033, USA. Tel.: +1 323 4421417.

E-mail address: jcamarer@usc.edu (J.A. Camarero).

1 Authors contributed equally to this work.

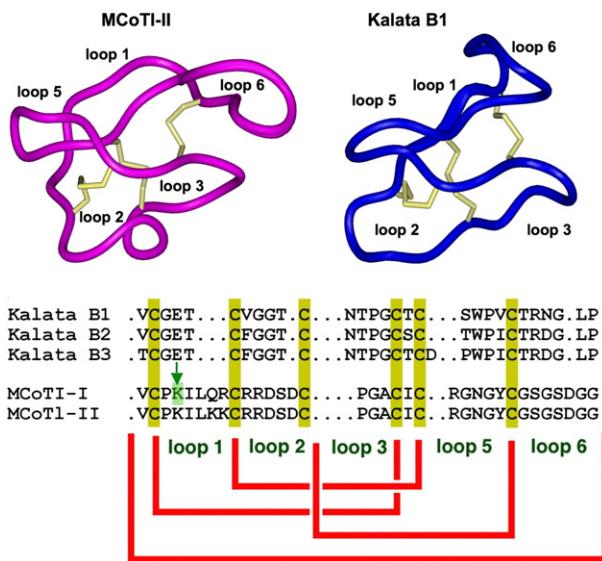


Fig. 1. Primary and tertiary structures of MCoTI and kalata cyclotides. The structures of MCoTI-II (pdb ID: 1IB9 [26]) and kalata B1 (pdb ID: 1NB1 [83]) are shown. Conserved cysteine residues and disulfide bonds are shown in yellow. An arrow marks residue Lys⁴ located at loop 1 in MCoTI-cyclotides. This residue in MCoTI-I was used for the site-specific conjugation of AlexaFluor488 N-hydroxysuccinimide ester (AF488-OSu) through a stable amide bond.

The latter method can generate folded cyclotides either *in vivo* or *in vitro* using standard bacterial expression systems [20–22] and opens the possibility of producing large libraries of genetically encoded cyclotides which can be analyzed by high throughput cell-based screening for selection of specific sequences able to bind to particular biomolecular targets [21,23].

Cyclotides have been classified into three main subfamilies. The Möbius and bracelet cyclotide subfamilies differ in the presence or absence of a *cis*-Pro residue, which introduces a twist in the circular backbone topology [24]. A third subfamily comprises the cyclic trypsin inhibitors MCoTI-I/II (Fig. 1), which have been recently isolated from the dormant seeds of *Momordica cochinchinensis*, a plant member of the *cucurbitaceae* family, and are powerful trypsin inhibitors ($K_i \approx 20$ –30 pM) [25]. These cyclotides do not share significant sequence homology with other cyclotides beyond the presence of the three-cystine bridges, but structural analysis by NMR has shown that they adopt a similar backbone-cyclic cystine-knot topology [26,27]. MCoTI cyclotides, however, show high sequence homology with related linear cystine-knot squash trypsin inhibitors [25], and therefore represent interesting molecular scaffolds for drug design [19,28–30]. Indeed, acyclic squash inhibitors have already been used as scaffolds for the incorporation of novel bioactive peptides to render de-novo engineered knottins with novel biological activities [31,32].

All these features make cyclotides ideal drug development tools [19,28–30]. They are remarkably stable due to the cyclic cystine knot [33]. They are relatively small, making them readily accessible to chemical synthesis [16]. They can also be encoded within standard cloning vectors, and expressed in cells [20–22], and are amenable to substantial sequence variation [34], making them ideal substrates for molecular grafting of biological peptide epitopes [4]. They are also amenable to molecular evolution strategies to enable generation and selection of compounds with optimal binding and inhibitory characteristics [22,34]. Even more importantly, MCoTI-cyclotides have been recently shown to be able to enter human macrophages and breast cancer cell lines [35]. Internalization into macrophages was shown to be mediated mainly through macropinocytosis, a

form of endocytosis that is actin-mediated and results in formation of large vesicles termed macropinosomes [36,37]. It should be noted, however, that in this study the visualization of MCoTI-II uptake was done in fixed, not live, cells. Analysis of live cells provides the ability to visualize events in real time without the possible complication of fixation artifacts that have confounded interpretations of the uptake of Tat and other related peptides for instance [38,39]. As well, macropinocytosis is a dominant mechanism for endocytic uptake in macrophages [40–42], unlike other cells that are not specialized for large-scale sampling of extracellular fluid, and which use multiple alternative endocytic mechanisms. These mechanisms can include clathrin-mediated endocytosis, caveolar endocytosis, macropinocytosis, phagocytosis, flotillin-dependent endocytosis, as well as multiple other as yet under-characterized mechanisms [43,44]. Intrigued by these results, we explored the cellular uptake of site-specific fluorescently-labeled MCoTI-I cyclotides and studied the cellular uptake mechanisms in HeLa cells using live cell imaging by confocal fluorescence microscopy.

In this work we report for the first time the cellular uptake of MCoTI-cyclotides monitored by real time confocal fluorescence microscopy imaging in live HeLa cells. Our results clearly show that HeLa cells readily internalize fluorescently-labeled-MCoTI-I. We found that this process is temperature-dependent and can be reversibly inhibited at 4 °C, which indicates an active mechanism of internalization. The internalized cyclotide also seems to colocalize in live cells with multiple endocytic markers including, to the greatest extent, the fluid-phase endocytic marker dextran (10KDa dextran, 10K-Dex). Internalized MCoTI-I was colocalized to a lesser extent with the cholesterol/lipid dependent endocytic marker cholera toxin B (CTX-B) and the clathrin-mediated endocytic marker, EGF. Internalized MCoTI-I was localized within a fairly rapid time course within late endosomal and lysosomal compartments which engaged in rapid and directed movements suggestive of cytoskeletal involvement. MCoTI-I uptake in HeLa cells was not impaired by Latrunculin B (Lat B) although its uptake was approximately 80% inhibited by the less specific macropinocytosis inhibitor EIPA. Altogether, these data seem to indicate that MCoTI-I cyclotide is capable of internalization in live cells through multiple endocytic pathways that may be dominant in the particular cell type under study and may also include macropinocytosis. The lack of strong preference for MCoTI-I internalization via a specific cellular internalization pathway is of significant value since the lack of endogenous affinity for a particular pathway can enable the ready re-targeting by introduction of targeting peptides, within the scaffold, that may enable specific and targeted endocytic uptake to a particular target cell. At the same time, the ready uptake of MCoTI-I by multiple pathways suggests accessibility, in the untargeted form, to essentially all cells.

2. Materials and methods

2.1. Analytical characterization of cyclotides

Analytical HPLC was performed on an HP1100 series instrument with 220 and 280 nm detection using a Vydac C18 column (5 µm, 4.6 × 150 mm) at a flow rate of 1 mL/min. Preparative and semi-preparative HPLC were performed on a Waters Delta Prep system fitted with a Waters 2487 UV-visible detector using a Vydac C18 (15–20 µm, 10 × 250 mm) at a flow rate of 5 mL/min. All runs used linear gradients of 0.1% aqueous trifluoroacetic acid (TFA, solvent A) vs. 0.1% TFA, 90% acetonitrile in H₂O (solvent B). Ultraviolet-visible (UV-vis) spectroscopy was carried out on an Agilent 8453 diode array spectrophotometer. Electrospray mass spectrometry (ES-MS) analysis was routinely applied to all compounds and components of reaction mixtures. ES-MS was performed on an Applied Biosystems API 3000 triple quadrupole electrospray mass spectrometer using

Analyst 1.4.2. Calculated masses were obtained using Analyst 1.4.2. All chemicals involved in synthesis or analysis were obtained from Aldrich (Milwaukee, WI) or Novabiochem (San Diego, CA) unless otherwise indicated.

2.2. Preparation of Fmoc-Tyr(tBu)-F

Fmoc-Tyr(tBu)-F was prepared using diethylaminosulfur trifluoride DAST [45] and quickly used afterwards. Briefly, to a stirred solution of Fmoc-Tyr(tBu)-OH (459.6 mg, 1 mmol) 10 mL of dry dichloromethane (DCM), containing dry pyridine (800 μ L, 1 mmol) and (1.1 mL, 1.2 mmol) DAST was added dropwise at 25 °C under nitrogen current. After 20 min, the mixture was washed with ice-cold water (3 \times 20 mL). The organic layer was separated and dried over anhydrous $MgSO_4$. The solvent was removed under reduced pressure to give the corresponding Fmoc-amino acyl fluoride as white solid that was used immediately. Amino acid fluorides should be used immediately as they are extremely unstable and prone to hydrolysis.

2.3. Loading of 4-sulfamylbutyryl AM resin with Fmoc-Tyr(tBu)-F

Loading of the first residue was accomplished using Fmoc-Tyr(tBu)-F according to standard protocols [46]. Briefly, 4-sulfamylbutyryl AM resin (420 mg, 0.33 mmol) (Novabiochem) was swollen for 20 min with dry DCM and then drained. A solution of Fmoc-Tyr(tBu)-F (\approx 461 mg, 1 mmol) in dry DCM (2 mL) and di-isopropylethylamine (DIEA) (180 μ L, 1 mmol) was added to the drained resin and reacted at 25 °C for 1 h. The resin was washed with dry DCM (5 \times 5 mL), dried and kept at –20 °C until use.

2.4. Chemical synthesis of MCoTI-I

Solid-phase synthesis was carried out on an automatic peptide synthesizer ABI433A (Applied Biosystems) using the Fast-Fmoc chemistry with 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) activation protocol at 0.1 mmol scale on a Fmoc-Tyr(tBu)-sulfamylbutyryl AM resin. Side-chain protection was employed as previously described for the synthesis of peptide α -thioesters by the Fmoc-protocol [47], except for the N-terminal Cys residue, which was introduced as Boc-Cys(Trt)-OH. After chain assembly, the alkylation, thiolytic cleavage and deprotection were performed as previously described [48,49]. Briefly, \approx 100 mg of protected peptide resin were first alkylated two times with ICH_2CN (174 μ L, 2.4 mmol; previously filtered through basic silica) and DIEA (82 μ L, 0.46 mmol) in N-methylpyrrolidone (NMP) (2.2 mL) for 12 h. The resin was then washed with NMP (3 \times 5 mL) and DCM (3 \times 5 mL). The alkylated peptide resin was cleaved with $HSCH_2CH_2CO_2Et$ (200 μ L, 1.8 mmol) in the presence of a catalytic amount of sodium thiophenolate ($NaSP_4$, 3 mg, 22 μ mol) in dimethylformamide (DMF):DCM (3:4 v/v, 1.4 mL) for 24 h. The resin was then dried at reduced pressure. The side-chain protecting groups were removed by treating the dried resin with trifluoroacetic acid (TFA): H_2O :tri-isopropylsilane (TIS) (95:3:2 v/v, 5 mL) for 3–4 h at room temperature. The resin was filtered and the linear peptide thioester was precipitated in cold Et_2O . The crude material was dissolved in the minimal amount of H_2O :MeCN (4:1) containing 0.1% TFA and characterized by HPLC and ES-MS as the desired MCoTI-I linear precursor α -thioester [expected mass (average isotopic composition) = 3608.2 Da; measured = 3608.8 \pm 0.3 Da]. Cyclization and folding was accomplished by flash dilution of the MCoTI-I linear α -thioester TFA crude to a final concentration of \approx 50 μ M into freshly degassed 2 mM reduced glutathione (GSH), 50 mM sodium phosphate buffer at pH 7.5 for 18 h. Folded MCoTI-I was purified by semi-preparative HPLC using a linear gradient of 10–35% solvent B over 30 min. Pure MCoTI-I was characterized by HPLC and ES-MS [expected mass (average isotopic composition) = 3480.9 Da; measured = 3481.0 \pm 0.4 Da].

2.5. Recombinant expression of MCoTI-I

Bacterial expression and purification of MCoTI-I was carried out as previously described [22].

2.6. Chemical labeling of MCoTI with AlexaFluor488 succinimide ester (AF488-NHS)

MCoTI-I was site-specifically labeled through the ε -amino of residue Lys⁴ (Fig. 1). MCoTI-I only has one Lys residue in its sequence (Fig. 1). Briefly, MCoTI-I (5 mg, 1.4 μ mol) was conjugated with two-fold molar excess of AF488-NHS in 0.2 M sodium phosphate buffer (2.5 mL) at pH 7.5 for 2 h. The reaction was quenched with 6 mM NH_2 -OH solution at pH 4. AF488-labeled MCoTI-I was purified by semi-preparative HPLC using a linear gradient of 15–35% solvent B over 30 min. Pure labeled MCoTI-I was characterized by HPLC and ES-MS [expected mass (average isotopic composition) = 3997.9 Da; measured = 3997.4 \pm 0.3 Da].

2.7. Purification of synthetic MCoTI-I using trypsin-Sepharose beads

Preparation of trypsin-Sepharose beads was done as previously described [21–23]. Pull down experiments with synthetic MCoTI-I were performed as follows: synthetic MCoTI-I cyclization/folding crude reactions were typically incubated with 0.2 mL of trypsin-Sepharose for one hour at room temperature with gentle rocking, and centrifuged at 3000 rpm for 1 min. The beads were washed with 50 volumes of PBS containing 0.1% Triton X-100, then rinsed with 50 volumes of PBS, and drained of excess PBS. Bound MCoTI-I was eluted with 0.4 mL of 8 M GdmCl and fractions were analyzed by RP-HPLC and ES-MS.

2.8. Endocytosis experiments

For studies of endocytic uptake mechanisms, methyl- β -cyclodextrin (MBCD) and 5-(N-ethyl-N-isopropyl)amiloride (EIPA) were purchased from Sigma-Aldrich. Latrunculin B (Lat B) was purchased from Calbiochem. Texas Red-EGF (TR-EGF), Lysotracker™ Red DN-99 (LysoRed), AF594 cholera toxin B (AF594 CTX-B), Texas Red 10,000 MW dextran (TR-10K Dex) and CellLight™ Lysosomes-RFP (RFP-Lamp1), rhodamine-phalloidin, and DAPI were all purchased from Invitrogen (Carlsbad, CA).

2.9. Cell culture

HeLa cells were obtained from the American Type Culture Collection (ATCC) and were cultured in a humidified incubator at 37 °C in 95% air/5% CO_2 in phenol red-free Dulbecco's modified essential medium (DMEM)(4.5 g/L glucose with 10% FBS, 1% glutamine, and 1% non-essential amino acids) and split with trypsin/EDTA as recommended by the manufacturer.

2.10. Confocal fluorescence microscopy

For MCoTI-I uptake studies, HeLa cells were seeded on 35 mm glass-bottom culture dishes (MatTek, Ashland, MA) at a density of 8.5×10^4 cells/dish. On day 2 of culture, the cells were rinsed with PBS and the media replaced with incubation buffer (phenol red-free, serum-free DMEM with 1% P/S and 20 mM HEPES) prior to addition of AF488-MCoTI-I (25 μ M) and incubation at 37 °C for 1 h. Following this time, excess MCoTI-I was rinsed off with a gentle PBS wash and the media replaced prior to imaging. Intracellular distribution was analyzed at 1 h and again at regular intervals for up to 10 h. For assessment of distribution and colocalization of AF488-MCoTI-I and Lysotracker™ Red, RFP-Lamp1, AF594-Cholera toxin B, TR-10K Dex, or TR-EGF in live cells, we utilized a Zeiss

LSM 510 Meta NLO imaging system equipped with Argon and HeNe lasers and mounted on a vibration-free table for confocal fluorescence microscopy. For analysis of the effects of Lat B or EIPA pre-treatment, 2 μ M Lat B or 50 μ M EIPA was added for 30 min at 37 °C prior to addition of AF488-MCoTI-I. For colocalization studies, Lysotracker™ Red (50 nM), AF594-Cholera toxin B (10 μ g/mL), TR-10K Dex (1 mg/mL), or TR-EGF (400 ng/mL) was added to cells simultaneously with MCoTI-I prior to incubation at 37 °C. Analysis of the extent of colocalization was done at 1 h of uptake. For colocalization with RFP-Lamp1, cells were treated with RFP-Lamp1-expressing BacMam (2×10^7 particles/plate) on the previous day. For temperature-dependent uptake studies, cells were cooled on ice for 30 min prior to the addition of AF488-MCoTI-I in incubation buffer. After incubation at 4 °C for 30 min, the cells were imaged and subsequently incubated at 37 °C for 1 h before imaging again. For fixation and visualization of actin filaments following treatment with or without 2 μ M Lat B, the cells were fixed with 4% paraformaldehyde prior to the addition of rhodamine-phalloidin and DAPI. For analysis of fluorescent pixel colocalization, cells from at least 3 different experiments were analyzed individually. Using the Zeiss LSM 510 software colocalization tool, regions of interest (ROI) were selected and marked with an overlay to encompass all pixels, following the Zeiss manual protocol. The threshold was automatically set from these ROIs. For time-lapse imaging, cells were incubated with 25 μ M AF488-MCoTI-I for 1 h at 37 °C. Following this time, excess MCoTI-I was rinsed off with a gentle PBS wash and the media replaced prior to imaging. The time series image capture was set to a 2.5 second delay between scans.

3. Results and discussion

In order to study the cellular uptake of MCoTI-cyclotides, we decided to use MCoTI-I. MCoTI-I contains only one Lys residue located in loop 1 versus MCoTI-II, which contains three Lys residues in the same loop (Fig. 1). The presence of only one Lys residue facilitates the site-specific introduction of a unique fluorophore on the sequence, thus minimizing any effect that the introduction of this group may have on the cellular uptake properties of the cyclotide.

Folded MCoTI-I cyclotide was produced either by recombinant or synthetic methods. In both cases the backbone cyclization was performed by an intramolecular native chemical ligation (NCL) [50–53] using the native Cys located on the beginning of loop 6 to facilitate the cyclization. This ligation site has been shown to give very good cyclization yields [21,22]. Intramolecular NCL requires the presence of an N-terminal Cys residue and C-terminal α -thioester group in the same linear precursor [52,54]. In the biosynthetic approach, the MCoTI-I linear precursor was fused in frame at the C- and N-terminus to a modified Mxe Gyrase A intein and a Met residue, respectively, and expressed in *Escherichia coli* [23]. This allows the generation of the required C-terminal thioester and N-terminal Cys residue after in vivo processing by endogenous Met aminopeptidase (MAP) [20,55]. Cyclization and folding can be accomplished very efficiently in vitro by incubating the MCoTI-I intein fusion construct in sodium phosphate buffer at pH 7.4 in the presence of reduced glutathione (GSH). Biosynthetic MCoTI-cyclotides generated this way have been shown to adopt a native folded structure by NMR and trypsin inhibitory assays [20,22,33].

Natively folded MCoTI-II has been already successfully produced using Fmoc-based solid-phase peptide synthesis [18,19]. Encouraged by these results we also explored the production of MCoTI-I by chemical synthesis (Fig. 2). For this purpose the MCoTI-I linear precursor α -thioester was assembled by Fmoc-based solid-phase peptide synthesis on a sulfonamide resin [48,49] (Fig. 2A). Activation of the sulfonamide linker with iodoacetonitrile, followed by cleavage with ethyl mercaptoacetate and acidolytic deprotection with TFA, provided the fully protected linear peptide α -thioester

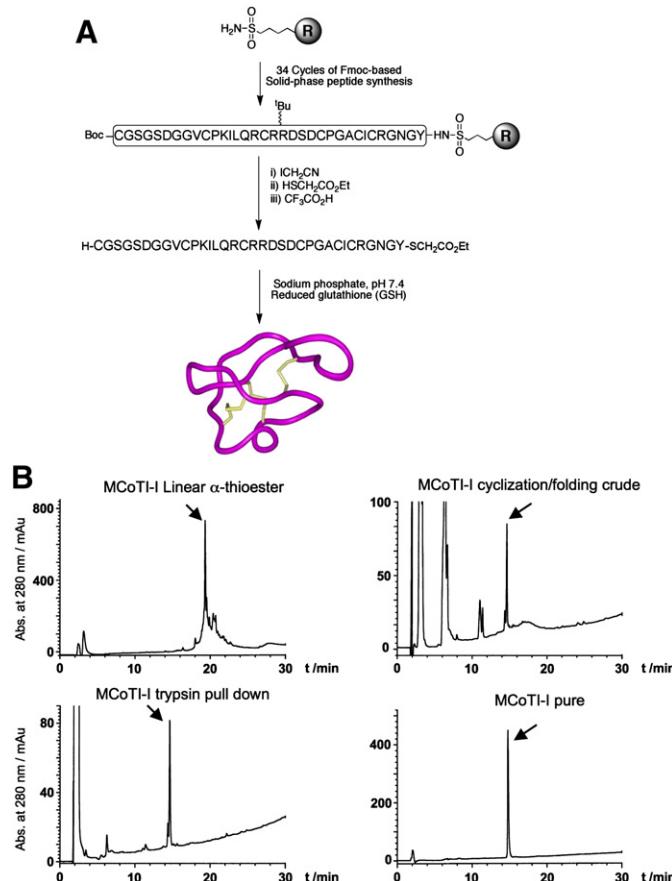


Fig. 2. Chemical synthesis of MCoTI-I. (A) Synthetic scheme used for the chemical synthesis of cyclotide MCoTI-I by Fmoc-based solid-phase peptide synthesis. (B) Analytical reverse-phase HPLC traces of MCoTI-I linear precursor α -thioester, cyclization/folding crude and purified MCoTI-I by either affinity chromatography using trypsin-immobilized Sepharose beads or semipreparative reverse-phase HPLC. HPLC analysis was performed in all the cases using a linear gradient of 0% to 70% buffer B over 30 min. Detection was carried out at 220 nm. An arrow indicates the desired product in each case.

(Fig. 2B). The synthetic linear precursor thioester was then efficiently cyclized and folded in a one-pot reaction using sodium phosphate buffer at pH 7.5 in the presence of 2 mM GSH. The reaction was complete in 18 h and the folded product was purified by reverse-phase HPLC and characterized by ES-MS. The expected mass for folded MCoTI-I was in agreement with a folded structure (expected mass = 3480.9 Da; measured = 3481.0 \pm 0.4 Da). Synthetic folded MCoTI-II was also shown to co-elute by HPLC with recombinant natively folded MCoTI-I (data not shown). The biological activity of synthetic MCoTI-I was assayed by using a trypsin pull-down experiment [22,23]. As shown in Fig. 2B, synthetic folded MCoTI-I was specifically captured from a cyclization/folding crude reaction by trypsin-immobilized Sepharose beads [21–23], thus indicating that it adopted a native cyclotide fold.

Purified MCoTI-I was site-specifically labeled with AlexaFluor488 (AF488) for live confocal imaging. The ε -amino group of Lys⁴ residue located in loop 1 was conjugated to AF488-NHS in sodium phosphate buffer at pH 7.5 for 2 h (Fig. 3A). Under these conditions, the main product of the reaction was mono-labeled AF488-MCoTI-I as characterized by HPLC and ES-MS (expected average mass = 3997.9 Da; measured = 3997.4 \pm 0.3 Da) (Fig. 3C). AF488-labeled MCoTI-I was then purified by reverse-phase HPLC to remove any trace of unreacted materials (Fig. 3B).

In order to infer the correct conclusions regarding data obtained on the cellular uptake of native MCoTI-I when using modified

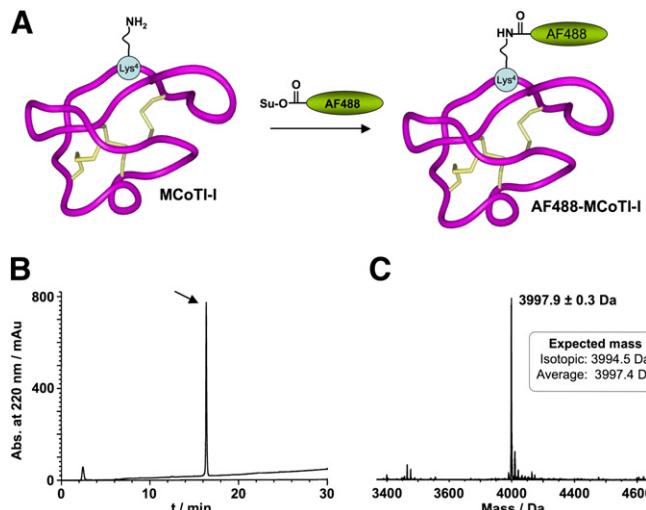


Fig. 3. Site-specific labeling of MCoTI-I with AlexaFluor488 N-hydroxysuccinimide ester (AF488-OSu). (A) Scheme depicting the bioconjugation process and localization of the fluorescent probe at residue Lys⁴ in loop 1. (B) Analytical reverse-phase HPLC trace of pure AF488-MCoTI-I. HPLC analysis was performed using a linear gradient of 0% to 70% buffer B over 30 min. Detection was carried at 220 nm. (C) ES-MS spectra of pure AF488-MCoTI-I.

cyclotides, like AF488-MCoTI-I for example, it is critical to be sure that they still adopt structures similar to that of the native form. MCoTI-cyclotides are extremely stable to chemical and thermal denaturation, and they have been shown to be able to withstand procedures like reverse-phase chromatography in the presence of organic solvents under acidic conditions without affecting their tertiary structure [16,18–21,33]. It is also unlikely that the acylation of the ϵ -amino group of Lys⁴ in MCoTI-I may disrupt the tertiary structure of this cyclotide. Craik and co-workers have previously shown that biotinylation of the three Lys residues located in loop 1 in MCoTI-II, including Lys⁴ (Fig. 1) does not disrupt the native cyclotide fold of this cyclotide as determined by ¹H-NMR [35]. We have also recently shown that mutation of residue Lys⁴ by Ala does not seem to affect the ability of this mutant to adopt a native cyclotide fold, thus indicating that the presence of positive charge residue in this position is not critical for the tertiary structure of MCoTI-I [22]. Similar findings have also been shown by Leatherbarrow and coworkers, where mutation of this residue by Phe or Val, rendered MCoTI-cyclotides able to fold correctly and have inhibitory activity against chymotrypsin and human elastase, respectively [19]. Altogether these facts suggest that residue Lys⁴ is not critical for adopting the native cyclotide fold or disturbing the tertiary structure of MCoTI-cyclotides.

To study the cellular uptake of AF488-MCoTI-I we used HeLa cells. The internalization studies were all carried out with 25 μ M AF488-MCoTI-I. This concentration provided a good signal/noise ratio for live cell confocal fluorescence microscopy studies and did not show any cytotoxic effect. This is in agreement with the cellular tolerance of wild-type and biotinylated MCoTI-II reported for other types of human cell lines [35]. First, we analyzed the time course of changes in cellular distribution following uptake of 25 μ M AF488-MCoTI-I by incubating with the cyclotide for 1 h and then analyzing its distribution after 1, 2, 4, 8 and 10 h. As shown in Fig. 4, the internalized cyclotide was clearly visible within perinuclear punctate spots inside the cells after 1 h incubation. Observation of cells pulsed with AF488-MCoTI-I for 1 h and then incubated for longer periods of time in the absence of cyclotide did not show any evidence for decreased intracellular fluorescence, while the largely perinuclear distribution of internalized MCoTI-I appeared comparable at all time points. Similar results have been also recently reported on the internalization

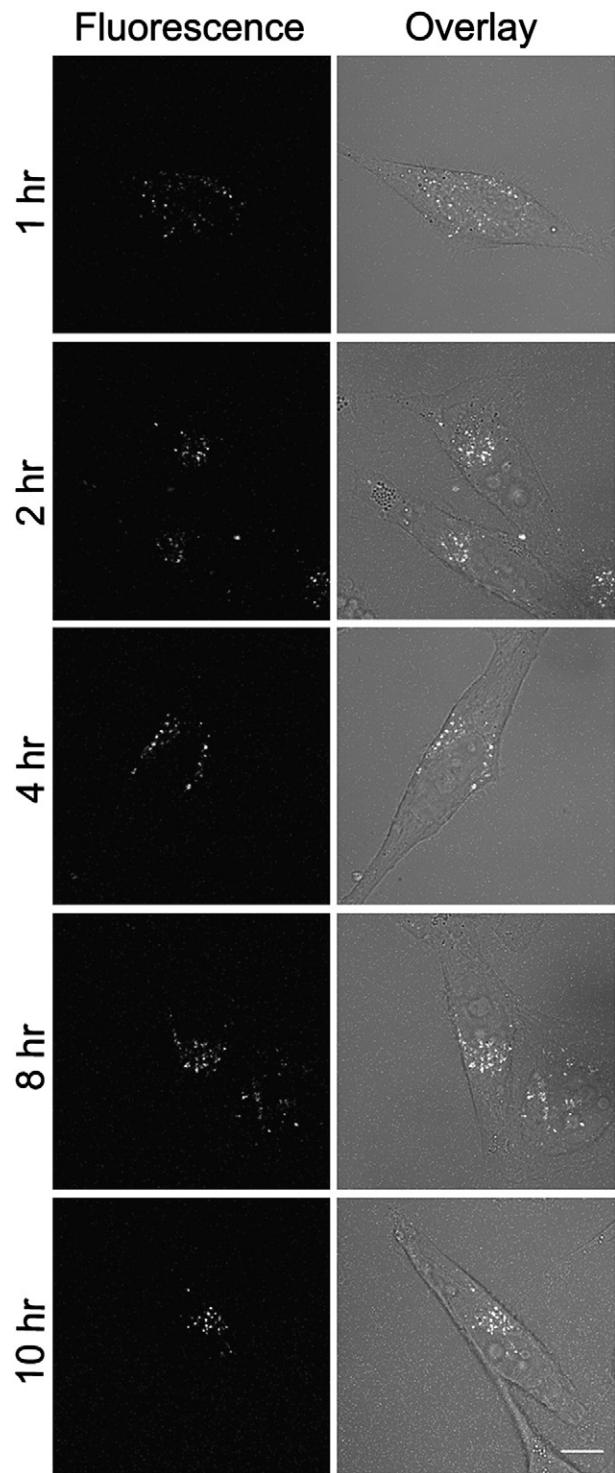


Fig. 4. MCoTI-I distribution in HeLa cells. HeLa cells were incubated with 25 μ M MCoTI-I for 1 h, followed by removal with gentle rinsing in PBS, and then monitored for distribution of intracellular fluorescence at intervals from 1 to 10 h using confocal fluorescence microscopy. Bar = 10 μ m.

of biotinylated MCoTI-II in macrophage and breast cancer cell lines [35], these studies however, used fixed cells to visualize the internalized cyclotide.

In order to study the mechanism of internalization of AF488-MCoTI-I in live HeLa cells, we first explored the effect of temperature on the uptake process. Active and energy-dependent endocytic mechanisms of internalization are inhibited at 4 °C [56]. The internalization

of AF488-MCoTI-I was totally inhibited after a 1 h incubation at 4 °C (Fig. 5). This inhibition was completely reversible and when the same cells were incubated again at 37 °C for 1 h, the punctate intracellular fluorescence labeling pattern was restored. This result confirmed that the uptake of AF488-MCoTI-I in HeLa cells follows a temperature dependent active endocytic internalization pathway. It should be noted that no significant surface binding was detected at 4 °C, suggesting that MCoTI-I does not bind a surface receptor, even nonspecifically. This is in agreement with studies on the MCoTI-II in fixed cells, so both the MCoTI-I and MCoTI-II appear to lack specific affinity for proteins or lipids in cell membranes, unlike the kalata B1 cyclotide which shows membrane affinity [35]. This lack of endogenous affinity for a specific surface receptor or membrane constituent makes MCoTI-I ideal for engineering using more specific, receptor-directed, peptide-based, internalization motifs, within the scaffold, that might enable members of this family to have enhanced targeting to a specific cell type.

Next, we investigated the internalization pathway used by labeled-MCoTI-I to enter HeLa cells. There are several known and well-characterized mechanisms of endocytosis [57]. It is also now well established that almost all cell-penetrating peptides (CPPs) use a combination of different endocytic pathways rather than a single endocytic mechanism [57]. A recent study showed that several CPPs (including Antennapedia/penetratin, nona-Arg and Tat peptides) can be internalized into cells by multiple endocytic pathways including macropinocytosis, clathrin-mediated endocytosis, and caveolae/lipid raft mediated endocytosis [58]. To investigate if that was the case with the internalization of AF488-MCoTI-I in HeLa cells, we decided to look at its colocalization with various endocytic markers (Fig. 6). 10K-Dex has previously been used as a marker of fluid-phase endocytosis [36,59–61]. CTX-B has been used as a marker for various lipid-dependent endocytic pathways [44,62], while EGF has traditionally been a marker of clathrin-mediated endocytosis [63–65]. As shown in Fig. 6, colocalization studies showed that after 1 h, AF488-MCoTI-I fluorescence was significantly colocalized with the fluorescence associated with 10K-Dex (59 ± 4% of total cyclotide fluorescence pixels were colocalized with 10K-Dex fluorescence pixels). Less colocalization was observed with fluorescent CTX-B

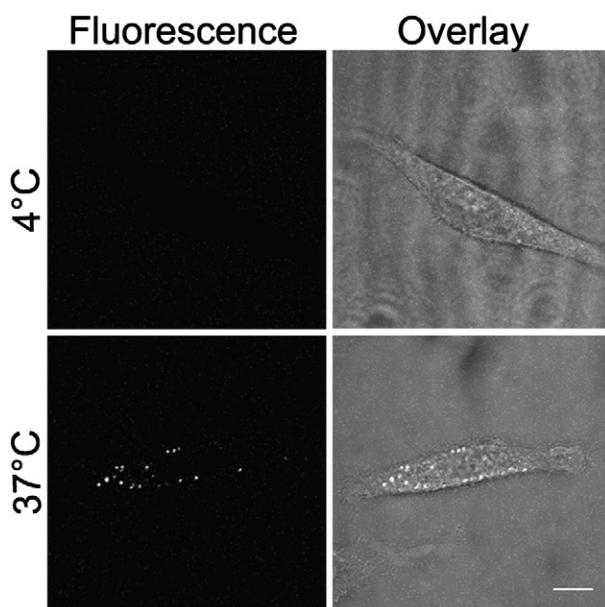


Fig. 5. Endocytosis of MCoTI-I is temperature-dependent. HeLa cells were incubated with 25 μ M MCoTI-I for 1 h at 4 °C. After removal of the MCoTI-I-containing media, and a gentle PBS wash, the cells were imaged. Following imaging, the MCoTI-I-containing media were replaced and the cells were incubated at 37 °C for 1 h and imaged again. Bar = 10 μ m.

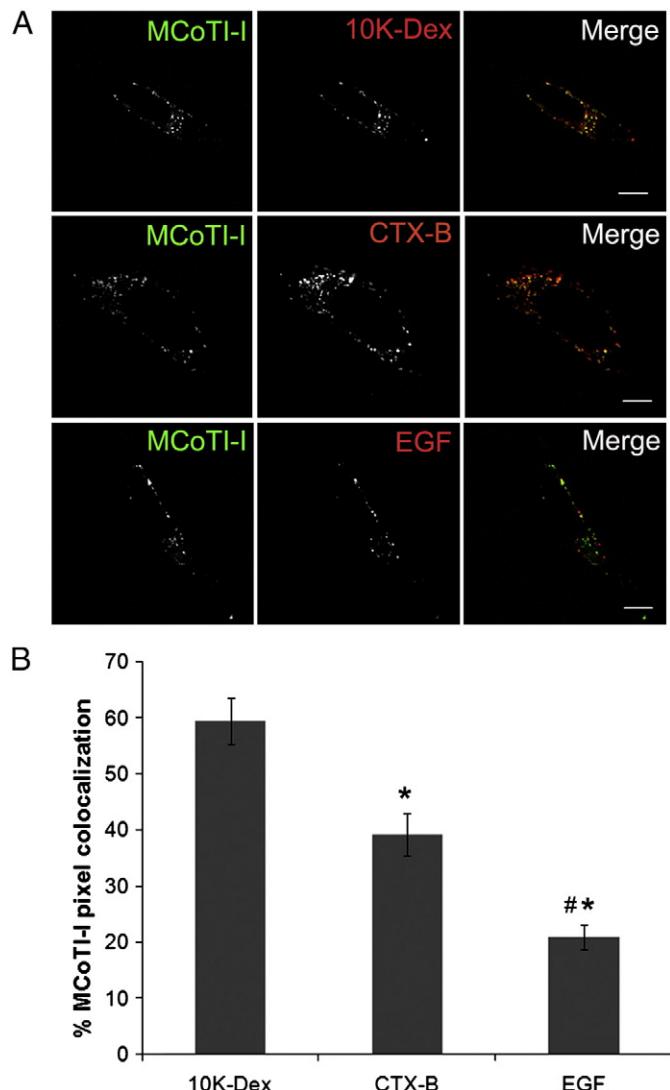


Fig. 6. Colocalization of MCoTI-I with markers of endocytosis. (A) HeLa cells were incubated with 25 μ M MCoTI-I and either 1 mg/mL 10K-dextran (10K-Dex), 10 μ g/mL cholera toxin B (CTX-B), or 400 ng/mL epidermal growth factor (EGF) for 1 h at 37 °C as described in Materials and methods and then imaged. Bar = 10 μ m. (B) Quantification of pixel colocalization was done using the Zeiss LSM software for image analysis and measures the % of total fluorescent AF488 MCoTI-I pixels in the ROI relative to red pixels associated with different endocytic markers. ($n = 13$ cells for 10K-Dex, $n = 11$ cells for CTX-B and $n = 10$ cells for EGF, with cells assessed across 3 different experiments, * $p \leq 0.05$ relative to 10K-Dex, # $p \leq 0.05$ relative to CTX-B).

(39 ± 4%) and fluorescent EGF (21 ± 2%). This data seems to suggest that AF488-MCoTI-I is primarily entering cells through fluid-phase endocytosis. The observed traces of colocalization with CTX-B and EGF also suggest that AF488-MCoTI-I could be using alternative or additional endocytic pathways. The colocalization results could also be attributed, however, to the merging of endosomal uptake vesicles generated by different pathways at the level of an early endosome. To address whether the major uptake and colocalization of AF488-MCoTI-I with 10K-Dex was due to cointernalization by macropinocytosis, we explored the inhibition of AF488-MCoTI-I uptake by Lat B, a potent inhibitor of actin polymerization, which is an essential element of macropinocytosis [66–69]. As shown in Fig. 7, Lat B did not significantly inhibit uptake of AF488-MCoTI-I (Fig. 7A) nor of 10K-Dex (data not shown). Treatment of HeLa cells with this agent resulted in a total disruption of the actin filament

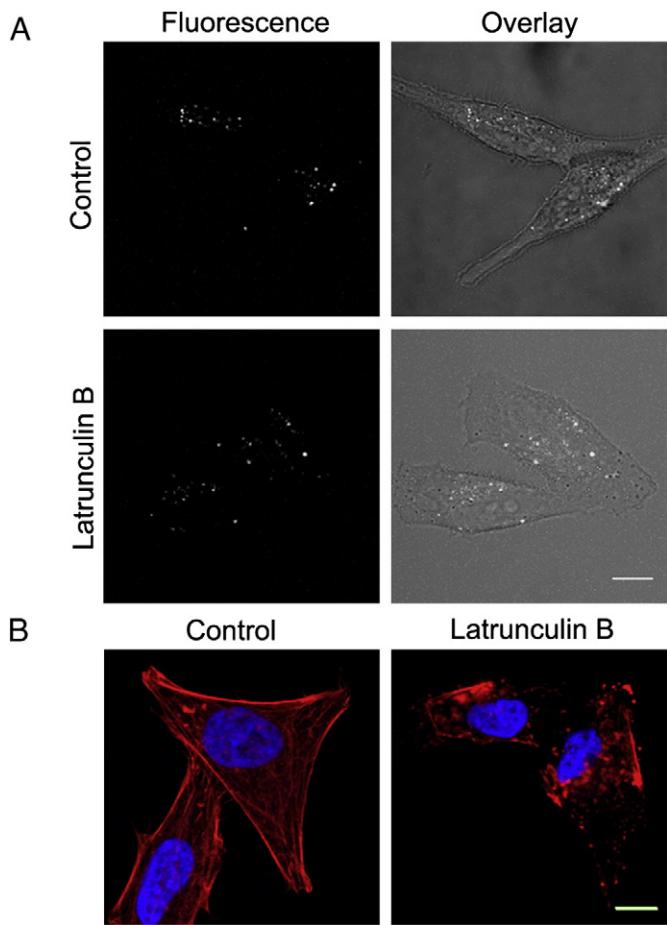


Fig. 7. Disruption of actin does not inhibit MCoTI-I uptake. (A) HeLa cells were untreated (control) or treated with Lat B (2 μ M) for 30 min at 37 °C prior to addition of 25 μ M MCoTI-I. Following uptake for 1 h at 37 °C, the cells were imaged using confocal fluorescence microscopy. Bar = 10 μ m. (B) HeLa cells without treatment (control) or treated with 2 μ M Lat B for 30 min at 37 °C were fixed and labeled with rhodamine-phalloidin to label actin (red) and DAPI to label nuclei (blue). Bar = 10 μ m.

network (Fig. 7B). We also explored the inhibition of AF488-MCoTI-I uptake by 5-N-ethyl-N-isopropyl amiloride (EIPA). EIPA is a potent and specific inhibitor of the Na^+/H^+ exchanger, whose activity is important for macropinosome formation [70]. Treatment of HeLa cells with 50 μ M EIPA significantly inhibited the uptake of both AF488-MCoTI-I ($\approx 80\%$) and 10K-Dex (Supplemental Fig. 1). The same treatment had no or little effect on the uptake of EGF (data not shown), but did significantly inhibit uptake of CTX-B (Supplemental Fig. 1), suggesting that at the concentration used, EIPA is affecting multiple endocytic mechanisms in these HeLa cells beyond simply macropinocytosis.

As an extension of these inhibition studies, cells were also treated with MBCD, a well-established cholesterol-depleting agent employed for studying the involvement of lipid rafts/caveolae in endocytosis [71,72]. Preliminary studies with MBCD suggested no significant inhibition of AF488-MCoTI-I (data not shown). Since the extent of total colocalization of AF488-MCoTI-I with CTX-B was less than 40%, it is unsurprising that no marked effect was seen by live cell microscopy. Taken together, these results seem to suggest that the uptake of AF488-MCoTI-I in HeLa cells is following multiple endocytic pathways, which is in agreement with what has been recently reported for different CPPs [58].

Next we explored the fate of the endocytic vesicles containing labeled MCoTI-I. There are at least two pathways that involve the

cellular trafficking of endosomal vesicles. The degradative pathway includes routing of internalized materials from early endosomes via late endosomes to lysosomes where degradation of internalized materials occurs within the cells. On the other hand, recycling endosomes sort material internalized into early endosomes and are responsible for effluxing internalized material back to the cellular membrane [73]. If labeled-MCoTI-I was localized in recycling endosomes, it would be expected that its concentration in the cell would decrease and/or accumulate on the membrane over time, which was not the case in the time course experiment following the cellular fate of internalized cyclotide (Fig. 4). To explore the potential localization of labeled-MCoTI-I in lysosomes we first used Lysotracker Red (LysoRed). This pH sensitive fluorescent probe is utilized for identifying acidic organelles, such as lysosomes and late endosomes, in live cells. As shown in Fig. 8A, significant colocalization ($60 \pm 4.0\%$ as determined by pixel colocalization analysis) of LysoRed and AF488-MCoTI-I was observed after treating the cells for 1 h with both agents. As an extension of these experiments, we also investigated the colocalization of labeled-MCoTI-I and lysosomal-associated membrane protein 1 (Lamp1), an established mature lysosomal marker [74,75]. For this experiment, live HeLa cells were first infected with a Red Fluorescent Protein (RFP)-Lamp1-expressing BacMam virus. The next day the cells were incubated with AF488-MCoTI-I for 1 h and imaged. As shown in Fig. 8B, colocalization was also seen for AF488-MCoTI-I and RFP-Lamp1 ($38 \pm 5\%$, as determined by pixel colocalization analysis), suggesting that even after 1 h, significant MCoTI-I has already reached the lysosomal compartments. Our data suggest that after 1 h, a significant amount of MCoTI-I ($\approx 40\%$) has trafficked through the endosomal pathway to the lysosomes and that $\approx 20\%$ is already localized in late endosomes or other types of acidic organelles. It has previously been reported that the perinuclear steady-state distribution of lysosomes is a balance between movement on microtubules and actin filaments [76–78]. Likewise, movement from early endosomal compartments to late endosomes to lysosomes has also been shown to rely on the microtubule network [79,80]. As an extension of these experiments, and to investigate whether MCoTI-I-containing vesicles were actively trafficking inside the cell, we captured time-lapse video of cells after incubation with MCoTI-I for 1 h. Indeed, the time-lapse capture showed active movements of MCoTI-I-containing vesicles (Fig. 9). Directed short- and long-range movements could be seen, characteristic of movement on cytoskeletal filaments. These results suggest that while a large portion of MCoTI-I has reached lysosomal compartments by 1 h, and some of the movements seen may be attributed to the steady-state distribution of lysosomes, the remaining cyclotide may still be trafficking through the cell from other membrane compartments, likely within late endosomes.

Craik and co-workers have recently reported the uptake of biotinylated-MCoTI-II by human macrophages and breast cancer cell lines [35]. This work concluded that the uptake of MCoTI-II in macrophages is mediated by macropinocytosis and that the cyclotide accumulates in macropinosomes without trafficking to the lysosome. MCoTI-II shares high homology with MCoTI-I ($\approx 97\%$ homology, see Fig. 1) and similar biological activity. Despite their similarities, the differences in the cellular uptake and trafficking of MCoTI-cyclotides by macrophages versus HeLa cells could be attributed to the cellular differences in endocytic preferences for these two very different cell types. Macrophage cells are specialized in large scale sampling of extracellular fluid using macropinocytosis as the dominant endocytic pathway. Meanwhile, other types of cells may employ multiple endocytic pathways as has been recently shown for the uptake of different CPPs in HeLa cells [58].

At this point we cannot be certain if some labeled MCoTI-I is able to escape from endosomal/lysosomal compartments into the cytosol. The ability to track the release of fluorescent-labeled molecules from

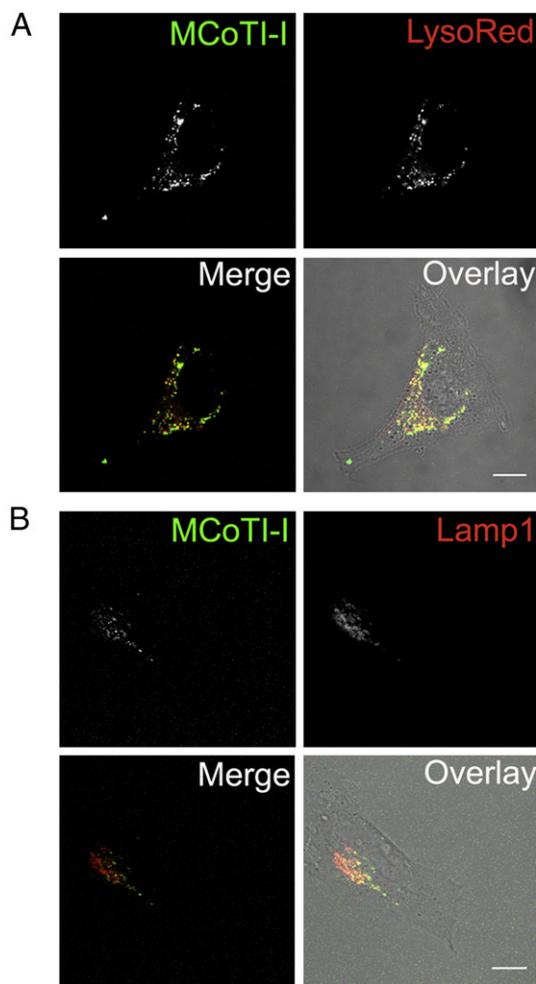


Fig. 8. MCoTI-I is colocalized with lysosomal compartments. Untreated (A) or BacMam-RFP-Lamp1 treated (B) HeLa cells were incubated with 25 μ M MCoTI-I and Lysotracker Red, or MCoTI-I alone, for 1 h at 37 °C as described in [Materials and methods](#) and then imaged. Bar = 10 μ m. (C) Quantification of % of total fluorescent AF488-MCoTI-I pixel colocalization with fluorescent pixels associated with both markers was done using the Zeiss LSM software for image analysis. ($n=14$ cells for Lysotracker Red and $n=11$ cells for Lamp1 with cells selected from 3 separate experiments).

cellular vesicles is limited using live cell imaging of fluorescence signal primarily due to the large dilution effect if the molecule is able to escape the highly confined volume of the vesicle into the larger cytosolic volume. One way to demonstrate the release of peptide into the cytosol, however, would be by using labels with better detection sensitivity or incorporating a biological activity that can be measured in the cellular cytosol. For example, the presence of Tyr residues in

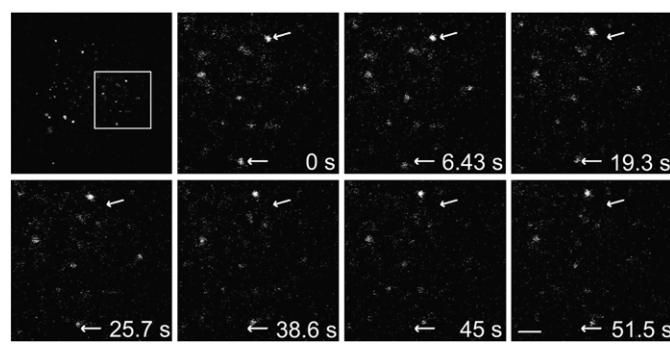


Fig. 9. MCoTI-I-containing vesicles are in motion. HeLa cells were incubated with 25 μ M MCoTI-I for 1 h at 37 °C and then imaged using time-lapse microscopy as described in [Materials and methods](#). Arrows indicate position of the moving vesicle at 0 min while displacement of the fluorescent vesicle relative to the arrow shows the extent of movement over time. Bar = 2 μ m.

both MCoTI-cyclotides should facilitate the incorporation of radioactive iodine into the phenolic ring of Tyr with minimal disruption of the native structure of the cyclotide. The incorporation or grafting of biological peptides into the MCoTI scaffold could also provide proof of endosomal/lysosomal escape if such biological activity could be measured only in the cytosol. This approach has already been used to demonstrate endosomal escape of CPPs such as the TAT peptide [81,82]. The retention of fluorescence signal in the perinuclear, lysosomal compartments for a period of up to 10 h suggests that most of the cyclotide remains within these compartments. However, given the flexibility of the cyclotide backbone to accommodate multiple peptide sequences, subsequent studies may explore the ability of targeting and endosomolytic sequences for concomitant targeted entry and endosomal/lysosomal escape into cytosol.

4. Conclusion

This study reports the first analysis of intracellular uptake of MCoTI-I cyclotide using live cell imaging by confocal fluorescence microscopy. Cyclotides represent a novel platform for drug development. Their stability, conferred by the cyclic cystine knot, their small size, their amenability to both chemical and biological synthesis, and their flexible tolerance to sequence variation make them ideal for grafting of biologically-active therapeutic epitopes. As we show herein, they are also capable, in the unmodified state, of utilizing multiple cellular endocytic pathways for internalization. Their ease of access makes them readily accessible, in their current state, to endosomal/lysosomal compartments of virtually any cell. Without an apparent strong preference for an existing cellular pathway, nor surface-expressed epitope in HeLa cells (nor in other studies with MCoTI-II in macrophages [35]), they appear highly amenable to retargeting to exploit a particular target cell's dominant internalization pathway and/or unique surface receptor repertoire, along with the targeted introduction of biologically-active therapeutic motifs.

Supplementary materials related to this article can be found online at [doi:10.1016/j.jconrel.2011.08.030](https://doi.org/10.1016/j.jconrel.2011.08.030).

Acknowledgments

The authors acknowledge the support of NIH grants EY017293 (to SHA), GM090323 (to JAC), Department of Defense Congressionally Directed Medical Research Program grant PC09305 (to JAC) and the support of a Ruth Kirchstein NIH minority predoctoral fellowship (F31EY018807) to JC.

References

[1] D.J. Craik, S. Simonsen, N.L. Daly, The cyclotides: novel macrocyclic peptides as scaffolds in drug design, *Curr. Opin. Drug Discovery Dev.* 5 (2) (2002) 251–260.

[2] D.J. Craik, M. Cemazar, C.H. Wang, N.L. Daly, The cyclotide family of circular miniproteins: nature's combinatorial peptide template, *Biopolymers* 84 (3) (2006) 250–266.

[3] K. Jagadish, J.A. Camarero, Cyclotides, a promising molecular scaffold for peptide-based therapeutics, *Biopolymers* 94 (5) (2010) 611–616.

[4] A.E. Garcia, J.A. Camarero, Biological activities of natural and engineered cyclotides, a novel molecular scaffold for peptide-based therapeutics, *Curr. Mol. Pharmacol.* 3 (3) (2010) 153–163.

[5] N.L. Daly, K.J. Rosengren, D.J. Craik, Discovery, structure and biological activities of cyclotides, *Adv. Drug Delivery Rev.* 61 (11) (2009) 918–930.

[6] O. Saether, D.J. Craik, I.D. Campbell, K. Sletten, J. Juul, D.G. Norman, Elucidation of the primary and three-dimensional structure of the uterotonin polypeptide kalata B1, *Biochemistry* 34 (13) (1995) 4147–4158.

[7] L. Cascales, D.J. Craik, Naturally occurring circular proteins: distribution, biosynthesis and evolution, *Org. Biomol. Chem.* 8 (22) (2010) 5035–5047.

[8] A.G. Poth, M.L. Colgrave, R.E. Lyons, N.L. Daly, D.J. Craik, Discovery of an unusual biosynthetic origin for circular proteins in legumes, *Proc. Natl. Acad. Sci. U. S. A.* 108 (25) (2011) 1027–1032.

[9] J.P. Mulvenna, C. Wang, D.J. Craik, CyBase: a database of cyclic protein sequence and structure, *Nucleic Acids Res.* 34 (2006) D192–D194 (Database issue).

[10] C.K. Wang, Q. Kaas, L. Chiche, D.J. Craik, CyBase: a database of cyclic protein sequences and structures, with applications in protein discovery and engineering, *Nucleic Acids Res.* 36 (2008) D206–D210 (Database issue).

[11] J.P. Mulvenna, J.S. Mylne, R. Bharathi, R.A. Burton, N.J. Shirley, G.B. Fincher, M.A. Anderson, D.J. Craik, Discovery of cyclotide-like protein sequences in graminaceous crop plants: ancestral precursors of circular proteins? *Plant Cell* 18 (9) (2006) 2134–2144.

[12] C.W. Gruber, A.G. Elliott, D.C. Ireland, P.G. Delprete, S. Dessein, U. Goransson, M. Trabi, C.K. Wang, A.B. Kinghorn, E. Robbrecht, D.J. Craik, Distribution and evolution of circular miniproteins in flowering plants, *Plant Cell* 20 (9) (2008) 2471–2483.

[13] I. Saska, A.D. Gillon, N. Hatsugai, R.G. Dietzgen, I. Hara-Nishimura, M.A. Anderson, D.J. Craik, An asparaginyl endopeptidase mediates in vivo protein backbone cyclization, *J. Biol. Chem.* 282 (40) (2007) 29721–29728.

[14] A.D. Gillon, I. Saska, C.V. Jennings, R.F. Guarino, D.J. Craik, M.A. Anderson, Biosynthesis of circular proteins in plants, *Plant J.* 53 (3) (2008) 505–515.

[15] B.F. Conlan, A.D. Gillon, D.J. Craik, M.A. Anderson, Circular proteins and mechanisms of cyclization, *Biopolymers* 94 (5) (2010) 573–583.

[16] N.L. Daly, S. Love, P.F. Alewood, D.J. Craik, Chemical synthesis and folding pathways of large cyclic polypeptides: studies of the cystine knot polypeptide kalata B1, *Biochemistry* 38 (32) (1999) 10606–10614.

[17] J.P. Tam, Y.A. Lu, J.L. Yang, K.W. Chiu, An unusual structural motif of antimicrobial peptides containing end-to-end macrocycle and cystine-knot disulfides, *Proc. Natl. Acad. Sci. U. S. A.* 96 (16) (1999) 8913–8918.

[18] P. Thongyoo, E.W. Tate, R.J. Leatherbarrow, Total synthesis of the macrocyclic cysteine knot microprotein MCoTI-II, *Chem. Commun. (Camb.)* 27 (2006) 2848–2850.

[19] P. Thongyoo, N. Roque-Rosell, R.J. Leatherbarrow, E.W. Tate, Chemical and biomimetic total syntheses of natural and engineered MCoTI cyclotides, *Org. Biomol. Chem.* 6 (8) (2008) 1462–1470.

[20] R.H. Kimura, A.T. Tran, J.A. Camarero, Biosynthesis of the cyclotide kalata B1 by using protein splicing, *Angew. Chem. Int. Ed.* 45 (6) (2006) 973–976.

[21] J.A. Camarero, R.H. Kimura, Y.H. Woo, A. Shekhtman, J. Cantor, Biosynthesis of a fully functional cyclotide inside living bacterial cells, *ChemBioChem* 8 (12) (2007) 1363–1366.

[22] J. Austin, W. Wang, S. Puttamadappa, A. Shekhtman, J.A. Camarero, Biosynthesis and biological screening of a genetically encoded library based on the cyclotide MCoTI-I, *ChemBioChem* 10 (16) (2009) 2663–2670.

[23] J. Austin, R.H. Kimura, Y.H. Woo, J.A. Camarero, In vivo biosynthesis of an Ala-scan library based on the cyclic peptide SFTI-1, *Amino Acids* 38 (5) (2010) 1313–1322.

[24] N.L. Daly, K.R. Gustafson, D.J. Craik, The role of the cyclic peptide backbone in the anti-HIV activity of the cyclotide kalata B1, *FEBS Lett.* 574 (1–3) (2004) 69–72.

[25] J.F. Hernandez, J. Gagnon, L. Chiche, T.M. Nguyen, J.P. Andrieu, A. Heitz, T. Trinh Hong, T.T. Pham, D. Le Nguyen, Squash trypsin inhibitors from *Momordica cochinchinensis* exhibit an atypical macrocyclic structure, *Biochemistry* 39 (19) (2000) 5722–5730.

[26] M.E. Felizmenio-Quimio, N.L. Daly, D.J. Craik, Circular proteins in plants: solution structure of a novel macrocyclic trypsin inhibitor from *Momordica cochinchinensis*, *J. Biol. Chem.* 276 (25) (2001) 22875–22882.

[27] A. Heitz, J.F. Hernandez, J. Gagnon, T.T. Hong, T.T. Pham, T.M. Nguyen, D. Le Nguyen, L. Chiche, Solution structure of the squash trypsin inhibitor MCoTI-II. A new family for cyclic knottins, *Biochemistry* 40 (27) (2001) 7973–7983.

[28] R.J. Clark, N.L. Daly, D.J. Craik, Structural plasticity of the cyclic-cystine-knot framework: implications for biological activity and drug design, *Biochem. J.* 394 (Pt. 1) (2006) 85–93.

[29] D.J. Craik, M. Cemazar, N.L. Daly, The cyclotides and related macrocyclic peptides as scaffolds in drug design, *Curr. Opin. Drug Discovery Dev.* 9 (2) (2006) 251–260.

[30] D.J. Craik, N.L. Daly, J. Mulvenna, M.R. Plan, M. Trabi, Discovery, structure and biological activities of the cyclotides, *Curr. Protein Pept. Sci.* 5 (5) (2004) 297–315.

[31] S. Reiss, M. Sieber, V. Oberle, A. Wentzel, P. Spangenberg, R. Claus, H. Kolmar, W. Losche, Inhibition of platelet aggregation by grafting RGD and KGD sequences on the structural scaffold of small disulfide-rich proteins, *Platelets* 17 (3) (2006) 153–157.

[32] H. Kolmar, Biological diversity and therapeutic potential of natural and engineered cystine knot miniproteins, *Curr. Opin. Pharmacol.* 9 (5) (2009) 608–614.

[33] S.S. Puttamadappa, K. Jagadish, A. Shekhtman, J.A. Camarero, Backbone dynamics of cyclotide MCoTI-I free and complexed with trypsin, *Angew. Chem. Int. Ed. Engl.* 49 (39) (2010) 7030–7034.

[34] H. Sanchez, J.A. Camarero, "Splicing up" drug discovery. Cell-based expression and screening of genetically-encoded libraries of backbone-cyclized polypeptides, *Adv. Drug Delivery Rev.* 61 (11) (2009) 908–917.

[35] K.P. Greenwood, N.L. Daly, D.L. Brown, J.L. Stow, D.J. Craik, The cyclic cystine knot miniprotein MCoTI-II is internalized into cells by macropinocytosis, *Int. J. Biochem. Cell Biol.* 39 (12) (2007) 2252–2264.

[36] L.J. Hewlett, A.R. Prescott, C. Watts, The coated pit and macropinocytic pathways serve distinct endosome populations, *J. Cell Biol.* 124 (5) (1994) 689–703.

[37] M.C. Kerr, R.D. Teasdale, Defining macropinocytosis, *Traffic* 10 (4) (2009) 364–371.

[38] M. Lundberg, M. Johansson, Positively charged DNA-binding proteins cause apparent cell membrane translocation, *Biochem. Biophys. Res. Commun.* 291 (2) (2002) 367–371.

[39] J.P. Richard, K. Melikov, E. Vives, C. Ramos, B. Verbeure, M.J. Gait, L.V. Chernomordik, B. Lebleu, Cell-penetrating peptides. A reevaluation of the mechanism of cellular uptake, *J. Biol. Chem.* 278 (1) (2003) 585–590.

[40] J.P. Lim, P.A. Gleeson, Macropinocytosis: an endocytic pathway for internalising large gulps, *Immunol. Cell Biol.* (2011).

[41] C.C. Norbury, L.J. Hewlett, A.R. Prescott, N. Shastri, C. Watts, Class I MHC presentation of exogenous soluble antigen via macropinocytosis in bone marrow macrophages, *Immunity* 3 (6) (1995) 783–791.

[42] R.M. Steinman, S.E. Brodie, Z.A. Cohn, Membrane flow during pinocytosis. A stereologic analysis, *J. Cell Biol.* 68 (3) (1976) 665–687.

[43] S.D. Conner, S.L. Schmid, Regulated portals of entry into the cell, *Nature* 422 (6927) (2003) 37–44.

[44] G.J. Doherty, H.T. McMahon, Mechanisms of endocytosis, *Annu. Rev. Biochem.* 78 (2009) 857–902.

[45] C. Kaduk, H. Wenschuh, M. Beyermann, K. Forner, L.A. Carpin, M. Bienert, Synthesis of Fmoc-amino acid fluorides via DAST, an alternative fluorinating agent, *Lett. Pept. Sci.* 2 (5) (1997) 285–288.

[46] R. Ingenito, D. Dreznjak, S. Guffler, H. Wenschuh, Efficient loading of sulfonamide safety-catch linkers by Fmoc amino acid fluorides, *Org. Lett.* 4 (7) (2002) 1187–1188.

[47] J.A. Camarero, A.R. Mitchell, Synthesis of proteins by native chemical ligation using Fmoc-based chemistry, *Protein Pept. Lett.* 12 (8) (2005) 723–728.

[48] R. Ingenito, E. Bianchi, D. Fattori, A. Pessi, Solid phase synthesis of peptide C-terminal thioesters by Fmoc/t-Bu chemistry Source, *J. Am. Chem. Soc.* 121 (49) (1999) 11369–11374.

[49] Y. Shin, K.A. Winans, B.J. Backes, S.B.H. Kent, J.A. Ellman, C.R. Bertozzi, Fmoc-based synthesis of peptide, *J. Am. Chem. Soc.* 121 (1999) 11684–11689.

[50] J.A. Camarero, T.W. Muir, α -Thioesters: application to the total chemical synthesis of a glycoprotein by native chemical ligation. Chemoslective backbone cyclization of unprotected peptides, *Chem. Commun.* 1997 (1997) 1369–1370.

[51] L. Zhang, J.P. Tam, Synthesis and application of unprotected cyclic peptides as building blocks for peptide dendrimers, *J. Am. Chem. Soc.* 119 (1997) 2363–2370.

[52] J.A. Camarero, J. Pavel, T.W. Muir, Chemical synthesis of a circular protein domain: evidence for folding-assisted cyclization, *Angew. Chem. Int. Ed.* 37 (3) (1998) 347–349.

[53] Y. Shao, W.Y. Lu, S.B.H. Kent, A novel method to synthesize cyclic peptides, *Tetrahedron Lett.* 39 (23) (1998) 3911–3914.

[54] J.A. Camarero, T.W. Muir, Biosynthesis of a head-to-tail cyclized protein with improved biological activity, *J. Am. Chem. Soc.* 121 (1999) 5597–5598.

[55] J.A. Camarero, D. Fushman, D. Cowburn, T.W. Muir, Peptide chemical ligation inside living cells: in vivo generation of a circular protein domain, *Bioorg. Med. Chem.* 9 (9) (2001) 2479–2484.

[56] S.C. Silverstein, R.M. Steinman, Z.A. Cohn, Endocytosis, *Annu. Rev. Biochem.* 46 (1977) 669–722.

[57] S.B. Fonseca, M.P. Pereira, S.O. Kelley, Recent advances in the use of cell-penetrating peptides for medical and biological applications, *Adv. Drug Delivery Rev.* 61 (11) (2009) 953–964.

[58] F. Duchtard, M. Fotin-Mleczek, H. Schwarz, R. Fischer, R. Brock, A comprehensive model for the cellular uptake of cationic cell-penetrating peptides, *Traffic* 8 (7) (2007) 848–866.

[59] D. Brown, I. Sabolic, Endosomal pathways for water channel and proton pump recycling in kidney epithelial cells, *J. Cell Sci. Suppl.* 17 (1993) 49–59.

[60] W. Shurety, N.L. Stewart, J.L. Stow, Fluid-phase markers in the basolateral endocytic pathway accumulate in response to the actin assembly-promoting drug Jasplakinolide, *Mol. Biol. Cell* 9 (4) (1998) 957–975.

[61] K. Thompson, M.J. Rogers, F.P. Coxon, J.C. Crockett, Cytosolic entry of bisphosphonate drugs requires acidification of vesicles after fluid-phase endocytosis, *Mol. Pharmacol.* 69 (5) (2006) 1624–1632.

[62] I.R. Nabi, P.U. Le, Caveolae/raft-dependent endocytosis, *J. Cell Biol.* 161 (4) (2003) 673–677.

[63] F. Huang, A. Khvorova, W. Marshall, A. Sorkin, Analysis of clathrin-mediated endocytosis of epidermal growth factor receptor by RNA interference, *J. Biol. Chem.* 279 (16) (2004) 16657–16661.

[64] X. Jiang, F. Huang, A. Marusyk, A. Sorkin, Grb2 regulates internalization of EGF receptors through clathrin-coated pits, *Mol. Biol. Cell* 14 (3) (2003) 858–870.

[65] A. Sorkin, M. Von Zastrow, Signal transduction and endocytosis: close encounters of many kinds, *Nat. Rev. Mol. Cell Biol.* 3 (8) (2002) 600–614.

- [66] C.J. Burckhardt, U.F. Greber, Virus movements on the plasma membrane support infection and transmission between cells, *PLoS Pathog.* 5 (11) (2009) e1000621.
- [67] G.V. Jerdeva, K. Wu, F.A. Yarber, C.J. Rhodes, D. Kalman, J.E. Schechter, S.F. Hamm-Alvarez, Actin and non-muscle myosin II facilitate apical exocytosis of tear proteins in rabbit lacrimal acinar epithelial cells, *J. Cell Sci.* 118 (Pt 20) (2005) 4797–4812.
- [68] J. Mercer, A. Helenius, Virus entry by macropinocytosis, *Nat. Cell Biol.* 11 (5) (2009) 510–520.
- [69] J.R. Peterson, T.J. Mitchison, Small molecules, big impact: a history of chemical inhibitors and the cytoskeleton, *Chem. Biol.* 9 (12) (2002) 1275–1285.
- [70] J. Mercer, A. Helenius, Vaccinia virus uses macropinocytosis and apoptotic mimicry to enter host cells, *Science* 320 (5875) (2008) 531–535.
- [71] P.U. Le, I.R. Nabi, Distinct caveolae-mediated endocytic pathways target the Golgi apparatus and the endoplasmic reticulum, *J. Cell Sci.* 116 (Pt 6) (2003) 1059–1071.
- [72] R.G. Parton, K. Simons, The multiple faces of caveolae, *Nat. Rev. Mol. Cell Biol.* 8 (3) (2007) 185–194.
- [73] A. Spang, On the fate of early endosomes, *Biol. Chem.* 390 (8) (2009) 753–759.
- [74] E.L. Eskelinen, Y. Tanaka, P. Saftig, At the acidic edge: emerging functions for lysosomal membrane proteins, *Trends Cell Biol.* 13 (3) (2003) 137–145.
- [75] M. Fukuda, Lysosomal membrane glycoproteins. Structure, biosynthesis, and intracellular trafficking, *J. Biol. Chem.* 266 (32) (1991) 21327–21330.
- [76] M.N. Cordonnier, D. Dauzon, D. Louvard, E. Coudrier, Actin filaments and myosin I alpha cooperate with microtubules for the movement of lysosomes, *Mol. Biol. Cell* 12 (12) (2001) 4013–4029.
- [77] R. Matteoni, T.E. Kreis, Translocation and clustering of endosomes and lysosomes depends on microtubules, *J. Cell Biol.* 105 (3) (1987) 1253–1265.
- [78] J. Taunton, B.A. Rowning, M.L. Coughlin, M. Wu, R.T. Moon, T.J. Mitchison, C.A. Larabell, Actin-dependent propulsion of endosomes and lysosomes by recruitment of N-WASP, *J. Cell Biol.* 148 (3) (2000) 519–530.
- [79] F. Aniento, N. Emans, G. Griffiths, J. Gruenberg, Cytoplasmic dynein-dependent vesicular transport from early to late endosomes, *J. Cell Biol.* 123 (6 Pt, 1) (1993) 1373–1387.
- [80] S. Loubry, C. Wilhelm, I. Hurbain, S. Neveu, D. Louvard, E. Coudrier, Different microtubule motors move early and late endocytic compartments, *Traffic* 9 (4) (2008) 492–509.
- [81] A.D. Frankel, C.O. Pabo, Cellular uptake of the tat protein from human immunodeficiency virus, *Cell* 55 (6) (1988) 1189–1193.
- [82] J.S. Lee, Q. Li, J.Y. Lee, S.H. Lee, J.H. Jeong, H.R. Lee, H. Chang, F.C. Zhou, S.J. Gao, C. Liang, J.U. Jung, FLIP-mediated autophagy regulation in cell death control, *Nat. Cell Biol.* 11 (11) (2009) 1355–1362.
- [83] K.J. Rosengren, N.L. Daly, M.R. Plan, C. Waine, D.J. Craik, Twists, knots, and rings in proteins. Structural definition of the cyclotide framework, *J. Biol. Chem.* 278 (10) (2003) 8606–8616.



NIH Public Access

Author Manuscript

Curr Pharm Des. Author manuscript; available in PMC 2012 April 20.

Published in final edited form as:

Curr Pharm Des. 2011 December ; 17(38): 4294–4307.

Cyclotides, a novel ultrastable polypeptide scaffold for drug discovery

Andrew Gould, Yanbin Ji, Teshome L. Aboye, and Julio A. Camarero*

Department of Pharmacology and Pharmaceutical Sciences, School of Pharmacy, University of Southern California, Los Angeles, CA 90033, USA

Abstract

Cyclotides are a unique and growing family of backbone cyclized peptides that also contain a cystine knot motif built from six conserved cysteine residues. This unique circular backbone topology and knotted arrangement of three disulfide bonds makes them exceptionally stable to thermal, chemical, and enzymatic degradation compared to other peptides of similar size. Aside from the conserved residues forming the cystine knot, cyclotides have been shown to have high variability in their sequence. Consisting of over 160 known members, cyclotides have many biological activities, ranging from anti-HIV, antimicrobial, hemolytic, and uterotonic capabilities; additionally, some cyclotides have been shown to have cell penetrating properties. Originally discovered and isolated from plants, cyclotides can be also produced synthetically and recombinantly. The high sequence variability, stability, and cell penetrating properties of cyclotides make them potential scaffolds to be used to graft known active peptides or engineer peptide-based drug design. The present review reports recent findings in the biological diversity and therapeutic potential of natural and engineered cyclotides.

Keywords

Cyclotides; cyclic peptides; peptide therapeutics; drug discovery; drug design

Introduction

Cyclotides are fascinating backbone-cyclized or circular proteins ranging from 28 to 37 amino acid residues that are naturally expressed in plants. They all share a unique head-to-tail circular knotted topology of three disulfide bridges, with one disulfide bond penetrating through a macrocycle formed by the other two disulfides bonds and interconnecting peptide backbones, forming what is called a cystine knot topology (Fig. 1) [1]. This cyclic cystine knot (CCK) framework gives cyclotides a compact, highly rigid structure [2], which confers exceptional resistance to thermal/chemical denaturation, and enzymatic degradation [3, 4]. In fact, the use of cyclotide-containing plants in indigenous medicine first highlighted the fact that these peptides are resistant to boiling and are apparently orally bioavailable [5].

More than 160 cyclotides have been isolated from plants in the Violaceae (violet), Rubiaceae (coffee) and Cucurbitaceae (squash) families [6, 7]; and more recently in the Fabaceae (legume) family [8–10]. It has been estimated, however, that around 50,000 cyclotides might exist [6]. The discovery of cyclotides in the plant *Clitoria ternatea* from the Fabaceae family represents an important discovery [9, 10]. The Fabaceae family is one the largest families of

*Address all correspondence to: Julio A. Camarero, Ph.D., Associate Professor, Department of Pharmacology and Pharmaceutical Sciences, University of Southern California, 1985 Zonal Avenue, PSC 616, Los Angeles, CA 90033, Tel: 323-442-1417, Fax: 323-224-7473, jcamarer@usc.edu.

plants on Earth, representing \approx 18,000 species, some of which are widely used as crops in human nutrition and food supply [11]. Due to the high diversity of is also expected that more cyclotides will be discovered in the near future.

Most of the cyclotides reported so far have been found in the Violaceae and Rubiaceae families. All the plants studied so far from the Violaceae family have shown to contain cyclotides. In contrast, only 5% of the Rubiaceae plants analyzed thus far have been shown to have cyclotides [6]. The only two cyclotides found to date from the Cucurbitaceae plant family are *Momordica cochinchinensis* trypsin inhibitor I and II (MCoTI-I/II; Fig. 1). These cyclotides are found in the seeds of *M. cochinchinensis* (a tropical squash plant) and are potent trypsin inhibitors. MCoTI cyclotides, however, do not share significant sequence homology with the other cyclotides beyond the presence of the three-cystine bridges that adopt a similar backbone-cyclic cystine-knot topology (Figs. 1 and 2). They are more related to linear cystine-knot squash trypsin inhibitors and sometimes are considered as circular knottins [7].

Despite the sequence diversity, all cyclotides share the same CCK motif (Figs.1 and 2). Hence, these micro-proteins can be considered natural combinatorial peptide libraries structurally constrained by the cystine-knot scaffold [12] and head-to-tail cyclization but in which hypermutation of most of the residues may be tolerated with the exception of the strictly conserved cysteines that comprise the knot.

All of the cyclotides reported so far from the Violaceae and Rubiaceae families are biosynthesized via processing from dedicated genes that in some cases encode multiple copies of the same cyclotide, and in others, mixtures of different cyclotide sequences [13]. Cyclotides from the Fabaceae family, however, are biosynthesized from evolved albumin-1 genes [9, 10].

Cyclotides can also be produced chemically using solid-phase peptide synthesis in combination with native chemical ligation [14–17] or recombinantly in bacteria by using modified protein splicing units or inteins [18–20]. The latter method can generate folded cyclotides either *in vivo* or *in vitro* using standard bacterial expression systems [18–20] and opens the possibility of producing large libraries of genetically encoded cyclotides which can be analyzed by high throughput cell-based screening for selection of specific sequences able to bind particular biomolecular targets [19, 20].

Naturally occurring cyclotides show various biological activities including insecticidal [15, 21, 22], uterotonic [23, 24], anti-viral [25, 26], antimicrobial [15, 24], antitumor [27], antihelminthic [28, 29] and protease inhibitory activity [30]. Their insecticidal and antihelminthic properties suggest that they may function as defense molecules in plants.

All these features make cyclotides ideal drug development tools [17, 31–33]. They are remarkably stable due to the cyclic cystine knot [2]. They are relatively small, making them readily accessible to chemical synthesis [14], and they can be encoded within standard cloning vectors and expressed in cells [18–20]. They are amenable to substantial sequence variation [34], which make them ideal substrates for molecular grafting of peptide epitopes [4] or for molecular evolution strategies to enable generation and selection of compounds with optimal binding and inhibitory characteristics [20, 34]. Even more importantly, some cyclotides have been shown recently to be able to cross cell membranes [35, 36].

In the present review, we discuss the structural and more pharmacologically-relevant biological properties of cyclotides as well as the latest developments in the use of the cyclotide scaffold to design novel peptide-based therapeutics.

The discovery

The discovery of the first cyclotide occurred indirectly when Norwegian researchers isolated and partially characterized a peptide, kalata B1, from the African plant *Oldenlandia affinis* [23, 37]. This plant was used by the natives to make a tea extract that was used to accelerate childbirth during labor, and thereafter kalata B1 was identified as the main uterotonic component of the plant [23, 37, 38]. Kalata B1 was initially characterized as a likely cyclic peptide of \approx 30 amino acids in size.

It was not until 1995, however, that the structure of kalata B1 was confirmed to be a 29-amino acid head-to-tail macrocyclic peptide containing a cystine knot motif by using a combination of mass spectrometry and NMR techniques [5]. The structural determination of kalata B1 coincided in time with the discovery of several other macrocyclic peptides of similar size and sequence from other Rubiaceae [39, 40] and Violaceae [41] plants. By 1999, another 20 different macrocyclic peptides featuring similar characteristics were described at the National Cancer Institute at USA [42] and by a Swedish group that was looking for novel cyclic peptides from plant biomass [43, 44]. In order to group this growing family of peptides, Craik and co-workers coined the term “cyclotides” (*cyclo peptides*) to refer to this interesting family of cyclic peptides. Since then, around 160 different cyclotides have been discovered in plants of the Rubiaceae, Violaceae, Cucurbitaceae and Fabaceae families and it has been estimated that the cyclotide family might contain over 50,000 different members [6, 45].

Cyclotides can be also classified in the larger family of knottins, a group of microproteins that are characterized for containing a cystine knot [46]. Several web-based databases have been created to curate the sequences of knottins (knottin.cbs.cnrs.fr) [47, 48] and cyclotides as well as other circular proteins (www.cybase.org.au) [49], both at the amino acid and nucleic acid levels.

Three-dimensional structure

Nearly all three dimensional structures of cyclotides elucidated to date have been performed by solution NMR spectroscopy [50]. The only exception is the cyclotide varv F, for which both an NMR and an X-ray crystal structure exists [51]. There are also several reports where the structures of cyclotides have been modeled [52, 53].

All of the natural cyclotides isolated thus far contain between 28 and 37 amino acids, including six cysteine residues (labeled Cys^I through Cys^{VI}), and are backbone cyclized (Figs. 1 and 2). The Cys residues are oxidized to form a cystine knot core in which the embedded ring formed by two disulfide bonds (Cys^I-Cys^{IV} and Cys^{II}-Cys^V) is penetrated by a third disulfide bond (Cys^{III}-Cys^{VI}). This unique topology of circular backbone and interlocked cystine core is usually referred to as a cyclic cystine knot (CCK) motif (Fig. 2) [54, 55]. The CCK motif is decorated by six interconnecting segments, or loops, which are successively numbered loop 1 through 6, starting at Cys^I (see Fig. 2). In most of the cyclotides, loops 1 and 4 are highly conserved in length and composition. In contrast, the other loops show more variation in size and sequence.

The first CCK structure was reported for the prototypic Möbius cyclotide kalata B1 in 1995 [5]. Since then NMR and X-ray studies have confirmed that all cyclotides retain a similar CCK scaffold but with high sequence variability in the six loops. The validity of the embedded knotted cystine core of cyclotides has been further confirmed by chemical means [56]. The CCK framework gives cyclotides a compact, highly rigid structure [2], which confers exceptional resistance to thermal/chemical denaturation and enzymatic degradation [3] thereby making cyclotides a promising molecular scaffold for drug discovery [1, 4, 57].

For example, cyclotides are more resistant to high temperatures approaching boiling when compared to their linear counterparts [3]. The stabilizing abilities of the CCK motif do not end there; the cystine knot also confers chemical stability to the protein allowing it to retain its native structure under typical denaturing conditions involving 8 M urea or 6 M guanidine hydrochloride [3]. Furthermore, the CCK motif also aids in protecting the proteins from enzymatic degradation by trypsin, pepsin, and thermolysin [3].

Cyclotides have been classified into three main subfamilies. The Möbius and bracelet cyclotide subfamilies differ in the presence or absence of a *cis*-Pro residue in loop 5, which introduces a twist in the circular backbone topology (Fig. 2B) [58]. Möbius and bracelet subfamilies of cyclotides may also be distinguished by the amount of hydrophobic residues and their location on the surface. For example, around 60% of the surface residues in the bracelet cyclotide cycloviolacin O2 are hydrophobic and are located primarily in loops 2 and 3. In contrast, only 40% of the surface residues in the Möbius cyclotide kalata B1 are hydrophobic and mainly localized in loops 2, 5 and 6 [59]. Nevertheless both major subfamilies of cyclotides contain some highly conserved residues such as Glu³ in loop 1 and the natural site of cyclization, Asn/Asp in loop 6. It is also interesting to remark that cyclotides that share properties of both Möbius and bracelet such as kalata B8 and psyle C are also starting to emerge [60, 61]. These cyclotides are usually referred to as hybrid cyclotides.

A third subfamily of cyclotides comprises the cyclic trypsin inhibitors MCoTI-I/II (Fig. 1), which have been isolated from the dormant seeds of *Momordica cochinchinensis*, a plant member of the Cucurbitaceae family, and are powerful trypsin inhibitors ($K_i \approx 20 - 30$ pM) [62]. These cyclotides do not share significant sequence homology with other cyclotides beyond the presence of the three-cystine bridges, but structural analysis by NMR has shown that they adopt a similar backbone-cyclic cystine-knot topology [63, 64].

Cyclotide plant biosynthesis

All the cyclotides reported so far from the Violaceae and Rubiaceae families are biosynthesized via processing from dedicated genes that, in some cases, encode multiple copies of the same cyclotide, and in others, mixtures of different cyclotide sequences [13, 21]. The first genes encoding cyclotides were discovered in the plant *Oldenlandia affinis* (Rubiaceae family) for the kalata cyclotides (Fig. 3A) [21]. For example, the gene encoding the precursor protein of kalata B1 (Oak 1) consists of an endoplasmic reticulum (ER)-targeting sequence, a pro-region, a highly conserved N-terminal repeat (NTR) region, a mature cyclotide domain, and a hydrophobic C-terminal tail (Fig. 3A). In contrast, the gene precursor of kalata B2 encodes up to three copies of the NTR-cyclotide region, which allows the formation of up to three mature kalata B2 cyclotides per molecule of precursor. Similar genes have been also found in plants from the Violaceae family [13, 45].

More recently, Craik and co-workers have revealed that the gene encoding the protein precursor of a novel cyclotide (Cter M) isolated from the leaf of butterfly pea (*Clitoria ternatea*), a representative member of the Fabaceae plant family, is embedded within the albumin-1 gene (Fig. 3A) [9, 10]. Similar findings have been also recently reported by Tam and co-workers [10]. Generic albumin-1 genes are comprised of an ER signal sequence followed by an albumin chain-b, a linker, and an albumin chain-a. In the precursor of cyclotide Cter M, the cyclotide domain replaces the albumin chain-b domain. The discovery that albumin genes can evolve into protein precursors that can be subsequently processed to become cyclic has also been described in a recent report on the biosynthesis of the peptide sunflower trypsin inhibitor 1 (SFTI-1) [65]. In this case the SFTI-1 precursor gene was identified as seed napin-like albumin [65]. SFTI-1 is a 14-residue peptide isolated from sunflower seeds with a head-to-tail cyclic backbone structure having only a single disulfide

bond. In this context, it is worth noting that the protein precursors of the only two cyclotides isolated so far from the *Cucurbitaceae* plant family (MCoTI-I/II) have not been identified yet.

The complete mechanism of how cyclotide precursors are processed and cyclized has not been completely elucidated yet. However, recent studies suggest that an asparaginyl endopeptidase (AEP) is a key element in the cyclization of cyclotides [66, 67]. It has been proposed that the cyclization step mediated by AEP takes place at the same time as the cleavage of the C-terminal pro-peptide from the cyclotide precursor protein through a transpeptidation reaction (Fig. 3B) [66]. The transpeptidation reaction involves an acyl-transfer step from the acyl-AEP intermediate to the N-terminal residue of the cyclotide domain [67]. AEPs are Cys proteases that are very common in plants and specifically cleave the peptide bond at the C-terminus of Asn and, less efficiently, Asp residues. All of the cyclotide precursors identified so far contain a well-conserved Asn/Asp residue at the C-terminus of the cyclotide domain in loop 6, which is consistent with the idea that cyclotides are cyclized by a transpeptidation reaction mediated by AEP (Fig. 3B) [66]. It is also worth noting that the only naturally linear 'cyclotide' isolated to date is violacin A from *Viola odorata* [68]. Interestingly, the gene encoding violacin A lacks the key Asn/Asp residue in loop 6, which has been replaced by a stop codon, and is required for the backbone cyclization reaction [68]. As a result, the peptide remains linear after translation and folding. Studies using transgenic plants expressing cyclotide precursors also support the involvement of AEP and requirement for a C-terminal Asn residue in the cyclotide sequence [66, 67]. For example, it has been shown that the introduction of a mutation at the C-terminal Asn residue of the cyclotide domain in transgenic plants resulted in no circular peptide production [67]. AEP has also been shown to be involved in the biosynthesis of the cyclic peptide SFTI-1 [65].

The expression of cyclotides in transgenic plants like *Arabidopsis* and tobacco is, however, highly inefficient giving rise to mostly acyclic or truncated proteins [66, 67], thus indicating that the processing and cyclization of cyclotides may involve other enzymes or may be species dependent. Craik and co-workers have recently isolated a protein disulfide isomerase (PDI) from the plant *O. affinis* showing that at least *in vitro* it is able to increase the folding yield of cyclotide-related molecules, including linear and cyclic cyclotides [69]. The relevance of this result *in vivo* still remains to be established.

Chemical Synthesis of cyclotides

Cyclotides are relatively small polypeptides, \approx 30–40 amino acids long, and therefore the linear precursors can be readily synthesized by chemical methods using solid-phase peptide synthesis (SPPS) [70]. Backbone cyclization of the corresponding linear precursor can readily be accomplished in aqueous buffers under physiological conditions by using an intramolecular version of native chemical ligation [71, 72] (Fig. 4). This approach has been successfully used to chemically generate native and engineered cyclotides [14, 16, 17, 73–76]. The only requirements for NCL are an N-terminal cysteine and an α -thioester group at the C-terminus of the linear precursor [77–80]. Both tert-butyloxycarbonyl (Boc)- and 9-fluorenyloxycarbonyl (Fmoc)-based chemistries have been used to incorporate C-terminal thioesters during chain assembly (Boc) [81–83] or using safety-catch based linkers (Fmoc) [80, 84–87]. Once the peptide is cleaved from the resin, both cyclization and folding can be carried out sequentially [14, 15, 74] or in a single pot reaction [16–18, 36].

Typically Möbius cyclotides such as kalata B1 are able to readily fold into its native structure quite efficiently [14] although during the oxidative folding a stable two-disulfide native-like intermediate accumulates. This intermediate is not the immediate precursor of the three-disulfide native peptide and is not observed during the reductive unfolding of

kalata B1 [88, 89]. MCoTI-cyclotides also fold very efficiently under oxidative conditions [16, 17, 36]. MCoTI-II also accumulates a native-like partially folded intermediate during its oxidative folding, but in contrast to kalata B1, this intermediate is a direct precursor to the fully folded native form [90].

The oxidative folding of bracelet cyclotides, on the other hand, has proven to be more challenging. For example, Tam and co-workers used orthogonal thiol protecting groups for the Cys residues in the synthesis of the bracelet cyclotides circulins A/B and cyclopyschotride [15, 74]. This approach allowed a more controlled step-wise folding strategy thereby improving the yield for the natively folded products. More recently, Göransson and co-workers have also shown that the bracelet cyclotide cycloviolacin O2 can be folded with yields over 50% using the mildly oxidizing agent dimethyl sulfoxide as co-solvent and a non-ionic detergent (Brij35) as a hydrophobic surface stabilizing agent [76].

Craik and co-workers have also recently reported that several analogs of the bracelet cyclotide cycloviolacin O1 resulting by the introduction of mutations on loops 2 and 6 can be efficiently folded [75]. The best analog in folding terms was accomplished by replacement of Ile in loop 2 with Gly together with the insertion of a Thr residue preceding Tyr in loop 6 [75]. These studies have allowed a better understanding of the structural elements influencing the efficiency of oxidative folding in Möbius and bracelet cyclotides thereby demonstrating that they are mostly localized within loops 2 and 6.

The oxidative folding of cyclotides is greatly influenced by the reaction conditions including the redox buffer used and the concentration of organic co-solvents. The Cys-containing tripeptide glutathione (GSH) is by far the most commonly employed reagent to accomplish the oxidative folding of cyclotides although other oxidizing reagents have also been used. Redox buffers containing GSH make possible the formation of disulfide bonds under thermodynamic control allowing disulfide reshuffling to recycle non-productive or misfolded intermediates [18, 76, 88, 89, 91]. GSH has been also used as a thiol additive as well as redox buffer for the cyclization and concomitant folding of cyclotides in a single pot reaction [18]. This has been successfully accomplished for Möbius [18, 20] and MCoTI [16, 17, 36] cyclotides (Fig. 3).

Leatherbarrow and coworkers have also reported the chemoenzymatic synthesis of several MCoTI cyclotides by using a protease-mediated ligation [17]. In this approach, the linear precursor was synthesized using Fmoc-SPPS, folded in oxidative conditions and cyclized using polymer-supported trypsin. This approach of enzyme-mediated *in vitro* cyclization has been also used for the cyclization of a linearized version of the cyclic peptide sunflower trypsin inhibitor, SFTI-1 [92]. These two studies suggest that protease-mediated cyclization can be a general and efficient process for producing cyclic peptides.

Recombinant expression of cyclotides

Our group has pioneered the recombinant production of cyclotides in bacteria through intramolecular NCL (see above) by using a modified protein-splicing unit or intein (Fig. 5) (see reference [34] for a recent review). Inteins are self-processing domains that undergo post-translational processing to splice together flanking external polypeptide domains (exteins) [93]. This approach uses a modified intein fused to the C-terminus of the cyclotide sequence to facilitate the formation of the required α -thioester at the C-terminus of the recombinant linear precursor (Fig. 5) [34]. The introduction of the required N-terminal Cys for cyclization can be easily accomplished by expressing the precursor protein with an N-terminal leading peptide, which can be cleaved by proteolysis or autoproteolysis (Fig. 5). The simplest way to achieve this is to introduce a Cys residue downstream to the initiating Met residue. Once the translation step is completed, the endogenous methionyl

aminopeptidases (MAP) removes the Met residue, thereby generating *in vivo* an N-terminal Cys residue [94–98]. The N-terminal Cys can then capture the reactive thioester in an intramolecular fashion to form a backbone cyclized polypeptide (Fig. 5).

Additional methods to generate N-terminal Cys proteases involve the use of exogenous proteases to cleave the leading signal after purification or *in vivo* by co-expressing the protease. Proteases that have been used so far include Factor Xa [79, 99], ubiquitin C-terminal hydrolase [100, 101], tobacco etch virus (TEV) protease [102], and thrombin [103]. The N-terminal pelB leader sequence has been also used recently to direct newly synthesized fusion proteins to the *E. coli* periplasmic space where the corresponding endogenous leader peptidases [104, 105] can generate the desired N-terminal cysteine-containing protein fragment [106]. Protein splicing can be also used to generate recombinant N-terminal Cys-containing polypeptides. For example, some inteins have been modified to allow cleavage at the C-terminal splice junction in a pH- and temperature-dependent fashion [107–109].

Intein-mediated backbone cyclization of polypeptides has also been recently used for the biosynthesis of the trypsin inhibitor SFTI-1 [110]. This method can also be applied to other Cys-rich peptides. The biosynthesis of backbone-cyclized α -defensins and naturally occurring cyclic θ -defensins is currently underway in our laboratory. It is worth noting that the recombinant expression of cyclotides facilitates the incorporation of NMR active isotopes such as ^{15}N and ^{13}C in a very inexpensive fashion, thus facilitating the use of the SAR by NMR (structure–activity relationship by nuclear magnetic resonance) [111] technique to characterize interactions between cyclotides and their biomolecular targets [19, 20] as well as carry out studies on the dynamics of the cyclotide scaffold [2]. The incorporation of ^{15}N into the cyclotide kalata B1 has been also recently reported by Craik and co-workers by whole plant labeling, i.e. by extracting the cyclotide from plants grown in the laboratory in media enriched in ^{15}N [112].

Biological activities of cyclotides

The natural function of cyclotides in plants seems to be as host-defense agents as deduced from their activity against insects [21, 22], nematodes [28, 29, 113, 114], and mollusks [60]. For example, it has been shown that cyclotides can efficiently inhibit the growth and development of insect and nematode larvae [115]. Although their mechanism of action is not totally well understood much of these activities seem to involve interaction of the cyclotide with membranes [9, 115]. Aside from their insecticidal and nematocidal activities, cyclotides have also been shown to have potential pharmacologically relevant activities, which include antimicrobial, anti-HIV, anti-tumor and neurotensin activities.

Antimicrobial activity—Most cyclotides have hydrophobic and hydrophilic patches located in different regions of their surface resembling to some extent the amphipathic character of classical antimicrobial peptides. The antimicrobial activities of cyclotides have been reported by two groups with conflicting results on the potency of kalata B1 against *Escherichia coli* and *Staphylococcus aureus*. In one study performed by Tam and co-workers, kalata B1 was active against *S. aureus*, but not *E. coli* [15], and in the second study, led by Gran and co-workers the peptide had the reverse effect [24]. This controversy has not been fully resolved yet, but it is most likely that different techniques were employed by both research teams. More recently, Tam and co-workers have also isolated several cyclotides from the plant *Clitoria ternatea* of the Fabaceae family with antimicrobial properties. In this work cyclotides CT1 and CT4 showed antimicrobial activity against strains of *E. coli*, *Klebsiella pneumonia* and *Pseudomonas aeruginosa* with minimal inhibitory concentrations ranging from 1 to 4 μM [10]. Craik and co-workers have also

identified similar cyclotides in the same plant [8, 9]. In this work, cyclotides CT1 and CT4 are referred as Cter P and Cter O, respectively. These bracelet cyclotides have high sequence homology with the cyclotide cycloviolacin O2, which has been shown to have antimicrobial activity against *P. aeruginosa* and a multiple drug resistant strain of *K. pneumonia* [116].

Antimicrobial peptides that target bacterial membranes typically have an amphipathic structure with a large number of positively charged residues, which explains their affinities for negatively charged membranes over more neutral mammalian cell membranes. Most cyclotides, however, have an overall charge close to zero at physiological pH, thereby making it unlikely that they interact with bacterial membranes through electrostatic interactions like classical cationic antimicrobial peptides [117]. Hence, further studies are required to investigate the antimicrobial properties and mechanism of action of cyclotides as well as its clinical relevance. This is particularly important given the growing occurrence of antibiotic resistance by microorganisms to current antibiotics.

Anti-HIV activity—The anti-HIV activity of cyclotides has been one of the most extensively studied so far due to its potential pharmacological applications [25, 39, 58, 118, 119]. The first account of a cyclotide with anti-HIV activity was reported by Gustafson and co-workers as part of a screening program at the U.S. National Cancer Center to search for novel natural products with anti-viral activity [39, 42]. In this work, several cyclotides isolated from the bark of the African tree *Chassalia parvifolia*, called circulins A-F, were shown to inhibit HIV infection in different host cell lines. Since then, several other cyclotides from the bracelet and Möbius subfamilies were also shown to have anti-HIV activity [25, 58, 118, 119].

The exact mode of action of these compounds remains still unclear although the inability of cyclotides to affect HIV reverse transcriptase activity combined with their cytoprotective effect suggests that the antiviral activity occurs before the entry of the virus into the host cell [39]. Recent studies have also suggested a strong correlation between the hydrophobic character of cyclotides and their anti-HIV activities [119, 120]. Moreover, surface plasma resonance and NMR studies have also shown that cyclotides can bind to model lipid membranes and that this interaction occurs primarily through their surface-exposed hydrophobic patches [121–123]. All these data suggest that the probable mode of anti-HIV action of cyclotides occurs through a mechanism that affects the binding and/or fusion of the virus to the target membrane. It is still premature, however, to conclude if cyclotide anti-HIV activity is the result of binding to the viral envelope, host cell membrane or both.

None of the cyclotides with anti-HIV activity, however, is being considered as candidates for anti-HIV therapy so far. This is due to their low therapeutic index (i.e. the ratio between the dose required for therapeutic effects versus toxic effects), which is typically too small (≤ 10) to be clinically useful. For example the therapeutic index of cyclotides kalata B1 and Varv E is 9 and 11, respectively. On the other hand, recent studies carried out on the cyclotides isolated from the plant *Viola yedoensis* (cycloviolacins Y4 and Y5) have shown therapeutic indexes of 45 and 14, respectively, which show some promise on the potential clinical use of these peptides to treat HIV infection [119].

Neurotensin antagonism—Cyclopyschotride (Cpt) A is a natural cyclotide isolated in 1994 from the South American tree *Psychotropia longipes* that has been reported to have neurotensin inhibition properties [40]. Neurotensin is a 13 amino acid neuropeptide that exerts its function by interacting with specific extracellular receptors increasing inositol triphosphate (IP₃) production and inducing Ca²⁺ mobilization from intracellular stores. Witherup and co-workers reported that Cpt A was able to inhibit neurotensin binding to its receptor in HT-29 cell membranes with an IC₅₀ ≈ 3 μ M [40]. The direct neurotensin

antagonism of Cpt A, however, was contradicted by the facts that it also increased intracellular Ca^{2+} levels, which could not be blocked by neurotensin antagonists [40]. Furthermore, Cpt A showed similar activity in two unrelated cell lines that did not express neurotensin receptors indicating that the mechanism of action is unlikely to be mediated through an interaction with the neurotensin receptor [40]. Recent studies using the cyclotide kalata B1 have shown that cyclotides are able to modulate membrane permeability by the formation of membrane pores with channel-like activity and no or little selectivity for specific cations [124]. The formation of similar channels by Cpt A could explain the increase on intracellular Ca^{2+} levels, although this has not been yet tested. Unfortunately there has been no follow-up on the neurotensin antagonism of Cpt A beyond the original report.

Antitumor activity—Several studies have reported the selective cytotoxicity of some cyclotides against cancer cells compared to normal cells [27, 125, 126]. In addition, the cytotoxic activity has been demonstrated using primary cancer cell lines [27]. Although the mechanism for cytotoxic activity of cyclotides is not totally well understood, it has been suggested that disturbance of the membrane integrity may be the cause for the cytotoxic activity. This is supported by the ability of cycloviolacin O2, one of the most active anticancer cyclotides, to disrupt tumor cell membranes [127]. Cancer cells differ from normal cells in the lipid and glycoprotein composition, which alters the overall net charge. For example, it has been shown that cancer cells express larger amounts of anionic phosphatidylserine phospholipids and O-glycosylated mucins, which typically confers a net negative charge to their membranes [128]. These differences are believed to play a major role in the cytotoxic selectivity of peptides with anticancer activity [128]. It is worth noting that cyclotide cytotoxicity is not only related to the three-dimensional structure but also to specific amino acid residues within the sequence [126, 129, 130]. For example, small modifications in the sequence of the bracelet cyclotide cycloviolacin O2 have been shown to have a great impact on cytotoxicity [129, 130]. Modification of the three positively charged residues in loops 5 and 6 as well as the Glu residue in loop 1 decreased the cytotoxic activity sevenfold and 48-fold, respectively [129]. More recent studies have also revealed that modification of the Trp residue in loop 2 has a detrimental effect on the cytotoxic activity of cycloviolacin O2 [130]. Similar results were also obtained for the cyclotide varv A, a Möbius cyclotide with high anticancer activity isolated from the plant *Viola arvensis* [130].

All of these studies suggest that differences in membrane composition and cyclotide primary sequence modulate membrane binding and the cytotoxic effects of cyclotides. A more comprehensive structural study on the membrane-binding properties of cyclotides will be required to have a better understanding on their antitumor mechanism of action.

Toxicity—Some cyclotides have been reported to have several toxic effects. For example, some cyclotides have been found to cause extensive hemolysis of human and rat erythrocytes, with HD_{50} (hemolytic dose) values ranging from 5 μM to 300 μM [15, 119, 131]. The large variation in HD_{50} value reflects different experimental conditions such as temperature and incubation time. The cyclotide kalata B1 has strong hemolytic activity and its lethal dose (LD_{50}) in rats and rabbits has been reported to be 1.0 mg/kg and 1.2 mg/kg, respectively, when it was injected intravenously. This cyclotide has been also reported to produce cardiotoxic effects associated with increased arterial blood pressure and tachycardia [38]. Interestingly, the strong hemolytic activity of kalata B1 can be eliminated by mutation to Ala of any one of eight residues located in the bioactive face of the molecule [132]. This suggests that it may be possible to eliminate other toxic effects by single mutation although this has yet to be tested.

It is worth noting that some other naturally occurring cyclotides have markedly reduced or no toxic effects at all. For example, the trypsin inhibitor cyclotides MCoTI-I and -II are non-hemolytic and non-toxic to human cells up to a concentration of 100 μ M [35], and therefore provide an excellent scaffold for the design of novel cyclotides with new biological activities. MCoTI-cyclotides are also particularly interesting from a pharmaceutical perspective because of their ability to penetrate cells, and therefore interact with intracellular targets [35, 36, 133].

Engineering cyclotides with novel activities

The unique properties associated with the cyclotide scaffold make them extremely valuable tools in drug discovery [1, 4, 117]. For example, the cyclic cystine knot (CCK) framework gives cyclotides a compact, highly rigid structure [2], which confers exceptional resistance to thermal/chemical denaturation, and enzymatic degradation [3]. Cyclotide can be also readily produced by chemical synthesis [14] and expressed in cells using standard cloning vectors [18–20]. Moreover, their high tolerance to mutations makes them ideal scaffolds for molecular grafting and evolution in order to generate novel cyclotides with new biological activities [20, 34]. Even more importantly, MCoTI-cyclotides have been shown recently to be able to enter human macrophages, breast and ovarian cancer cell lines [35, 36].

The high plasticity and tolerance to substitution of the cyclotide scaffold was first demonstrated by replacing some of the hydrophobic residues in loop 5 of cyclotide kalata B1 with polar and charged residues [31]. The mutated cyclotides retained the native fold of kalata B1, but were no longer hemolytic [31]. More recently our group has also performed a similar study on the cyclotide MCoTI-I, where all residues located in loops 1 through 5 were replaced by different amino acids. In this study only 2 from a total of 26 mutants were not able to fold efficiently [20]. These studies demonstrate the high plasticity of the cyclotide scaffold to mutations thus opening the possibility to introduce or graft foreign sequences onto them without affecting the native fold. Figure 6 highlights several studies that have used the cyclotide molecular scaffold to graft peptide sequences and to generate libraries for the purpose of engineering cyclotides with novel biological functions. It should be noted however, that the great stability and robustness of cyclotide framework makes necessary to be careful when grafting a peptide sequence into the peptide scaffold. It is important to be sure that the structure of the peptide displayed on the cyclotide will not be distorted by the conformation requirements of the cyclotide leading to a non-biologically active conformation of the grafted sequence.

The potential pharmaceutical applications of grafted cyclotides was first demonstrated in two recent studies aimed to develop novel anti-cancer [75] and anti-viral peptide-based therapeutics [17] (Fig. 6). The development of anti-cancer cyclotides involved grafting a peptide antagonist of angiogenesis onto the kalata B1 scaffold [75]. Tumor growth is usually associated with unregulated angiogenesis and therefore molecules with anti-angiogenic activity have potential applications in cancer treatment. In this study an Arg-rich peptide antagonist for the interaction of vascular endothelial growth factor A (VEGF-A) and its receptor was individually grafted into loops 2, 3 and 5 of kalata B1 [75]. The cyclotide grafted into loop 3 showed the highest activity in blocking VEGF-A receptor binding ($IC_{50} \approx 12 \mu$ M) when compared with that of the isolated grafted peptide epitope as well as the other grafted cyclotides. Although this is the first example of a successful functional redesign of a cyclotide, it should be noted that the biological activity would still need to be improved by several orders of magnitude for a potential application *in vivo*.

Additionally, the hemolytic activity of the grafted kalata B1 was completely removed and the proteolytic susceptibility of the Arg-rich grafted peptide was greatly diminished when compared to that of the isolated peptide sequence. This study clearly demonstrates that

cyclotides can be used as molecular scaffolds for displaying and stabilizing pharmacologically relevant active peptide sequences.

The utility of the cyclotide scaffold in drug design has also been recently shown by engineering non-native activities into the cyclotide MCoTI-II [17]. MCoTI-II is a member of the trypsin inhibitor subfamily. Trypsin inhibitor cyclotides are strong inhibitors of trypsin ($K_i \approx 20$ pM) and other trypsin-like proteases. In these cyclotides loop 1 is responsible for binding to the selectivity pocket of trypsin [62]. In this work, the loop 1 of MCoTI-II was engineered to be able to inhibit other proteases [17]. Interestingly, several analogs showed activity against the foot-and-mouth-disease virus (FMDV) 3C protease in the low μ M range [17]. The FMDV 3C protease is a Cys protease key for viral replication and therefore a potential target for the development of novel anti-viral therapeutics. This is the first reported peptide-based inhibitor for this protease. Although the potency was relatively low, this study demonstrates the potential of using MCoTI-based cyclotides for designing novel protease inhibitors [17].

In a more recent study, the same authors were also able to generate inhibitors of the serine proteases β -tryptase and human leukocyte elastase (HLE) using the cyclotide MCoTI-II as a molecular scaffold [134]. These two proteases are pharmacologically relevant drug targets that have been associated with respiratory and pulmonary disorders (HLE) and implicated in allergic and inflammatory disorders (β -tryptase). In this work Leatherbarrow and co-workers replaced the P1 residue in loop 1 to produce several MCoTI-II mutants (K6A and K6V) that showed activity against HLE in the low nM range ($K_i \approx 25$ nM) and relatively low activity against trypsin ($K_i \geq 1$ μ M) [134]. Interestingly, the same authors also showed that removal of the SDGG peptide segment in loop 6 yielded a relative potent β -tryptase inhibitor ($K_i \approx 10$ nM) without significantly altering the three-dimensional structure of the resulting cyclotide as determined by NMR [134]. The authors hypothesized that the removal of the Asp residue from loop 6 removes repulsive electrostatic and steric interactions with β -tryptase thus improving the inhibitory constant 160-fold when compared to the wild-type MCoTI-II.

Kolmar and co-workers have also recently reported the design of a series of inhibitors of human mast cell tryptase beta using the cyclotide MCoTI-II as scaffold [135]. In this interesting work, the authors introduced additional positive charge in the loop 6 of MCoTI-II. The resulting engineered cyclotides were able to inhibit all the monomers of the tryptase beta tetramer with K_i values around 1 nM.

Proteases are well-recognized drug targets and many diseases including inflammatory and pulmonary diseases, cancer, cardiovascular and neurodegenerative conditions have been associated with abnormal expression levels of proteases [136]. These two examples demonstrate that trypsin inhibitor cyclotides can be re-engineered to tailor their specificity for proteases other than trypsin, which has potential applications in drug development for protease targeting. The availability

Screening of cyclotide-based libraries

The ability to produce natively folded cyclotides *in vivo* discussed earlier opens up the intriguing possibility of generating large libraries of genetically-encoded cyclotides potentially containing billions of members. This tremendous molecular diversity should allow the selection strategies mimicking the evolutionary processes found in nature to select novel cyclotide sequences able to target specific molecular targets.

The potential for generating cyclotide libraries was first explored by our group using the kalata B1 scaffold [18]. In this work wild-type and several mutants of kalata B1 were

biosynthesized using an intramolecular native chemical ligation facilitated by a modified protein splicing unit (Fig. 5). The study generated six different linear versions of kalata B1, which were expressed in *E. coli* as fusions to a modified version of the yeast vacuolar membrane ATPase (VMA) intein. The *in vitro* folding and cyclization of the different kalata B1 linear precursors did not occur equally, suggesting the predisposition to adopt a native fold of the corresponding linear precursor may determine the efficiency of the cleavage/cyclization step [18]. This information was used to express a small library based on the kalata B1 scaffold. Cyclization and concomitant folding of the library was successfully performed *in vitro* using a redox buffer containing reduced GSH as a thiol co-factor therefore mimicking the intracellular conditions [18].

We have also demonstrated that this approach can be used for the production of cyclotides inside live *E. coli* cells [19]. In this study, the cyclotide MCoTI-II was efficiently produced in living *E. coli* cells by *in vivo* processing of the corresponding intein fusion precursor. In order to improve the expression yield of the precursor protein and boost the expression of folded cyclotide the bacterial gyrase A intein from *Mycobacterium xenopus* was used in this study instead [19]. This intein has been shown to express at higher yields than the yeast VMA intein in *E. coli* expression systems [19]. Using this approach, folded MCoTI-I can be expressed in *E. coli* cells to an intracellular concentration of low μM [19].

More recently, our group has extended this technology to the biosynthesis of a genetically encoded library of MCoTI-I based cyclotides [20]. In this case, the cyclization/folding of the library was performed either *in vitro*, by incubation with a redox buffer containing GSH, or by *in vivo* self-processing of the corresponding cyclotide-intein precursors. The cyclotide library was purified and screened for activity using trypsin-immobilized sepharose beads. The library was designed to mutate every single amino acid in loops 1, 2, 3, 4 and 5 to explore the effects on folding and trypsin binding activity of the resulting mutants. Interestingly, only two mutations (G27P and I22G) out of the 26 substitutions studied were able to negatively affect the folding of the resulting cyclotides. The I22G mutation affects loop 4, which is only formed by one residue. This loop forms part of the cyclotide scaffold, which may explain the deficient folding of this mutant. The G27P mutation is located at the end of loop 5. Intriguingly, this position is occupied by a Pro residue in Möbius cyclotides and is required for efficient folding. It is also interesting to remark that although these two mutants were not able to fold efficiently, the natively folded form was still able to bind trypsin [20]. The rest of the mutants were able to cyclize and fold with similar yields. As expected, the mutant K6A-MCoTI-I was not able to bind trypsin under the conditions used in the experiment although adopted a native cyclotide fold as determined by NMR [20]. As mentioned before, this residue is key for binding to the specificity pocket of trypsin, and it can be modified to change the inhibitory specificity of the resulting MCoTI-cyclotide to target other proteases [17, 134]. The affinity of each member of the MCoTI-library for trypsin was assayed using a competitive trypsin-binding assay [20]. The mutant cyclotides with less affinity were mostly found in loop 1 and the C-terminal region of loop 6, both well conserved among other squash trypsin inhibitors. These results combined with similar studies performed in kalata B1 [132] confirm the high plasticity and tolerance to mutations of the cyclotide framework, thus providing an ideal scaffold for the biosynthesis of large combinatorial libraries inside living bacterial cells. These genetically-encoded libraries can be screened *in-cell* for biological activity using high-throughput flow cytometry techniques for the rapid selection of novel biologically active cyclotides [18, 137, 138].

Concluding remarks

Cyclotides are a new emerging family of highly stable plant-derived backbone-cyclized polypeptides that share a disulfide-stabilized core characterized by an unusual knotted structure. Their unique circular backbone topology and knotted arrangement of three

disulfide bonds provides a compact, highly rigid structure [2] which confers exceptional resistance to thermal/chemical denaturation, and enzymatic degradation [3]. They can be also chemically synthesized allowing the introduction of chemical modifications such as non-natural amino acids and PEGylation to improve their pharmacological properties [14, 16, 17, 139]. Cyclotides can also be expressed in bacterial cells, and are amenable to substantial sequence variation, thus making them ideal substrates for molecular evolution strategies to enable generation and selection of compounds with optimal binding and inhibitory characteristics [18–20]. Finally, cyclotides have been shown to be able to cross human cell membranes [35, 36, 133]. Folded cyclotides are extremely resistant to chemical, physical and proteolytic degradation [3, 4]. Cyclotides have been also shown to fold inside bacterial cells [19, 20], which have a more reductive cytosolic environment than eukaryotic cells and therefore is highly unlikely that they were reduced in the cytosol of mammalian cells. Altogether, these characteristics make them promising leads or frameworks for peptide drug design [4, 31, 32, 117].

Acknowledgments

This work was supported by National Institutes of Health Research Grant R01-GM090323 and Department of Defense Congressionally Directed Medical Research Program Grant PC09305.

References

1. Daly NL, Rosengren KJ, Craik DJ. Discovery, structure and biological activities of cyclotides. *Adv Drug Deliv Rev.* 2009; 61:918–930. [PubMed: 19470399]
2. Puttamadappa SS, Jagadish K, Shekhtman A, Camarero JA. Backbone Dynamics of Cyclotide MCoTI-I Free and Complexed with Trypsin. *Angew Chem Int Ed Engl.* 2010; 49:7030–7034. [PubMed: 20715250]
3. Colgrave ML, Craik DJ. Thermal, chemical, and enzymatic stability of the cyclotide kalata B1: the importance of the cyclic cystine knot. *Biochemistry.* 2004; 43:5965–5975. [PubMed: 15147180]
4. Garcia AE, Camarero JA. Biological activities of natural and engineered cyclotides, a novel molecular scaffold for peptide-based therapeutics. *Curr Mol Pharmacol.* 2010; 3:153–163. [PubMed: 20858197]
5. Saether O, Craik DJ, Campbell ID, Sletten K, Juul J, Norman DG. Elucidation of the primary and three-dimensional structure of the uterotonic polypeptide kalata B1. *Biochemistry.* 1995; 34:4147–4158. [PubMed: 7703226]
6. Gruber CW, Elliott AG, Ireland DC, Delprete PG, Dessein S, Goransson U, et al. Distribution and evolution of circular miniproteins in flowering plants. *Plant Cell.* 2008; 20:2471–2483. [PubMed: 18827180]
7. Chiche L, Heitz A, Gelly JC, Gracy J, Chau PT, Ha PT, et al. Squash inhibitors: from structural motifs to macrocyclic knottins. *Curr Protein Pept Sci.* 2004; 5:341–349. [PubMed: 15551519]
8. Poth AG, Colgrave ML, Philip R, Kerenga B, Daly NL, Anderson MA, et al. Discovery of cyclotides in the fabaceae plant family provides new insights into the cyclization, evolution, and distribution of circular proteins. *ACS chemical biology.* 2010; 6:345–355. [PubMed: 21194241]
9. Poth AG, Colgrave ML, Lyons RE, Daly NL, Craik DJ. Discovery of an unusual biosynthetic origin for circular proteins in legumes. *Proc Natl Acad Sci U S A.* 2011; 108:1027–1032.
10. Nguyen GK, Zhang S, Nguyen NT, Nguyen PQ, Chiu MS, Hardjojo A, et al. Discovery and characterization of novel cyclotides originated from chimeric precursors consisting of albumin-1 chain a and cyclotide domains in the fabaceae family. *J Biol Chem.* 2011
11. Camarero JA. Legume cyclotides shed light on the genetic origin of knotted circular proteins. *Proc Natl Acad Sci U S A.* 2011; 108:10025–10026. [PubMed: 21653883]
12. Craik DJ, Cemazar M, Wang CK, Daly NL. The cyclotide family of circular miniproteins: nature's combinatorial peptide template. *Biopolymers.* 2006; 84:250–266. [PubMed: 16440288]

13. Dutton JL, Renda RF, Waine C, Clark RJ, Daly NL, Jennings CV, et al. Conserved structural and sequence elements implicated in the processing of gene-encoded circular proteins. *J Biol Chem.* 2004; 279:46858–46867. [PubMed: 15328347]
14. Daly NL, Love S, Alewood PF, Craik DJ. Chemical synthesis and folding pathways of large cyclic polypeptides: studies of the cystine knot polypeptide kalata B1. *Biochemistry.* 1999; 38:10606–10614. [PubMed: 10441158]
15. Tam JP, Lu YA, Yang JL, Chiu KW. An unusual structural motif of antimicrobial peptides containing end-to-end macrocycle and cystine-knot disulfides. *Proc Natl Acad Sci U S A.* 1999; 96:8913–8918. [PubMed: 10430870]
16. Thongyoo P, Tate EW, Leatherbarrow RJ. Total synthesis of the macrocyclic cysteine knot microprotein MCoTI-II. *Chem Commun (Camb).* 2006:2848–2850. [PubMed: 17007393]
17. Thongyoo P, Roque-Rosell N, Leatherbarrow RJ, Tate EW. Chemical and biomimetic total syntheses of natural and engineered MCoTI cyclotides. *Org Biomol Chem.* 2008; 6:1462–1470. [PubMed: 18385853]
18. Kimura RH, Tran AT, Camarero JA. Biosynthesis of the cyclotide kalata B1 by using protein splicing. *Angew Chem Int Ed.* 2006; 45:973–976.
19. Camarero JA, Kimura RH, Woo YH, Shekhtman A, Cantor J. Biosynthesis of a fully functional cyclotide inside living bacterial cells. *Chembiochem.* 2007; 8:1363–1366. [PubMed: 17590879]
20. Austin J, Wang W, Puttamadappa S, Shekhtman A, Camarero JA. Biosynthesis and biological screening of a genetically encoded library based on the cyclotide MCoTI-I. *Chembiochem.* 2009; 10:2663–2670. [PubMed: 19780078]
21. Jennings C, West J, Waine C, Craik D, Anderson M. Biosynthesis and insecticidal properties of plant cyclotides: the cyclic knotted proteins from *Oldenlandia affinis*. *Proc Natl Acad Sci U S A.* 2001; 98:10614–10619. [PubMed: 11535828]
22. Jennings CV, Rosengren KJ, Daly NL, Plan M, Stevens J, Scanlon MJ, et al. Isolation, solution structure, and insecticidal activity of kalata B2, a circular protein with a twist: do Möbius strips exist in nature? *Biochemistry.* 2005; 44:851–860. [PubMed: 15654741]
23. Gran L. Isolation of oxytocic peptides from *Oldenlandia affinis* by solvent extraction of tetraphenylborate complexes and chromatography on Sephadex LH-20. *Lloydia.* 1973; 36:207–208. [PubMed: 4744557]
24. Gran L, Sletten K, Skjeldal L. Cyclic Peptides from *Oldenlandia affinis* DC. Molecular and Biological Properties. *Chem Biodivers.* 2008; 5:2014–2022. [PubMed: 18972522]
25. Gustafson KR, McKee TC, Bokesch HR. Anti-HIV cyclotides. *Curr Protein Pept Sci.* 2004; 5:331–340. [PubMed: 15544529]
26. Hallock YF, Sowder RC, Pannell LK, Hughes CB, Johnson DG, Gulakowski R, et al. Cycloviolins A-D, anti-HIV macrocyclic peptides from *Leonia cymosa*. *J Org Chem.* 2000; 65:124–128. [PubMed: 10813905]
27. Lindholm P, Göransson U, Johansson S, Claeson P, Gullbo J, Larsson R, et al. Cyclotides: A novel type of cytotoxic agents. *Mol Cancer Ther.* 2002; 1:365–369. [PubMed: 12477048]
28. Colgrave ML, Kotze AC, Ireland DC, Wang CK, Craik DJ. The anthelmintic activity of the cyclotides: natural variants with enhanced activity. *Chembiochem.* 2008; 9:1939–1945. [PubMed: 18618891]
29. Colgrave ML, Kotze AC, Kopp S, McCarthy JS, Coleman GT, Craik DJ. Anthelmintic activity of cyclotides: In vitro studies with canine and human hookworms. *Acta Trop.* 2009; 109:163–166. [PubMed: 19059189]
30. Avrutina O, Schmoldt HU, Gabrijelcic-Geiger D, Le Nguyen D, Sommerhoff CP, Diederichsen U, et al. Trypsin inhibition by macrocyclic and open-chain variants of the squash inhibitor MCoTI-II. *Biol Chem.* 2005; 386:1301–1306. [PubMed: 16336125]
31. Clark RJ, Daly NL, Craik DJ. Structural plasticity of the cyclic-cystine-knot framework: implications for biological activity and drug design. *The Biochemical journal.* 2006; 394:85–93. [PubMed: 16300479]
32. Craik DJ, Cemazar M, Daly NL. The cyclotides and related macrocyclic peptides as scaffolds in drug design. *Current opinion in drug discovery & development.* 2006; 9:251–260.

33. Craik DJ, Daly NL, Mulvenna J, Plan MR, Trabi M. Discovery, structure and biological activities of the cyclotides. *Curr Protein Pept Sci.* 2004; 5:297–315. [PubMed: 15544527]

34. Sancheti H, Camarero JA. “Splicing up” drug discovery. Cell-based expression and screening of genetically-encoded libraries of backbone-cyclized polypeptides. *Adv Drug Deliv Rev.* 2009; 61:908–917. [PubMed: 19628015]

35. Greenwood KP, Daly NL, Brown DL, Stow JL, Craik DJ. The cyclic cystine knot miniprotein MCoTI-II is internalized into cells by macropinocytosis. *Int J Biochem Cell Biol.* 2007; 39:2252–2264. [PubMed: 17693122]

36. Contreras J, Elnagar AYO, Hamm-Alvarez S, Camarero JA. Cellular Uptake of cyclotide MCoTI-I follows multiple endocytic pathways. *J Control Release.* 201110.1016/j.jconrel.2011.1008.1030

37. Gran L. Oxytocic principles of Oldenlandia affinis. *Lloydia.* 1973; 36:174–178. [PubMed: 4744554]

38. Gran L. On the effect of a polypeptide isolated from “Kalata-Kalata” (Oldenlandia affinis DC) on the oestrogen dominated uterus. *Acta Pharmacol Toxicol (Copenh).* 1973; 33:400–408. [PubMed: 4801084]

39. Gustafson KR, Sowder RC, Louis E, Henderson LE, Parsons IC, Kashman Y, Cardellina JH, et al. Circulins A and B. Novel human immunodeficiency virus (HIV)-inhibitory macrocyclic peptides from the tropical tree Chassalia parvifolia. *J Am Chem Soc.* 1994; 116:9337–9338.

40. Witherup KM, Bogusky MJ, Anderson PS, Ramjit H, Ransom RW, Wood T, et al. Cyclopsyhotrope A, a biologically active, 31-residue cyclic peptide isolated from Psychotria longipes. *J Nat Prod.* 1994; 57:1619–1625. [PubMed: 7714530]

41. Schöpke T, Hasan Agha MI, Kraft R, Otto A, Hiller K. Hämolytisch aktive Komponenten aus *Viola tricolor* L. und *Viola arvensis* Murray. *Sci Pharm.* 1993; 61:145–153.

42. Gustafson KR, Walton LK, Sowder RC Jr, Johnson DG, Pannell LK, Cardellina JH Jr, et al. New circulin macrocyclic polypeptides from Chassalia parvifolia. *J Nat Prod.* 2000; 63:176–178. [PubMed: 10691702]

43. Claeson P, Goransson U, Johansson S, Luijendijk T, Bohlin L. Fractionation Protocol for the Isolation of Polypeptides from Plant Biomass. *J Nat Prod.* 1998; 61:77–81. [PubMed: 9548831]

44. Goransson U, Luijendijk T, Johansson S, Bohlin L, Claeson P. Seven novel macrocyclic polypeptides from *Viola arvensis*. *J Nat Prod.* 1999; 62:283–286. [PubMed: 10075760]

45. Simonsen SM, Sando L, Ireland DC, Colgrave ML, Bharathi R, Goransson U, et al. A continent of plant defense peptide diversity: cyclotides in Australian Hybanthus (Violaceae). *Plant Cell.* 2005; 17:3176–3189. [PubMed: 16199617]

46. Kolmar H, Skerra A. Alternative binding proteins get mature: rivalling antibodies. *FEBS J.* 2008; 275:2667. [PubMed: 18435760]

47. Gelly JC, Gracy J, Kaas Q, Le-Nguyen D, Heitz A, Chiche L. The KNOTTIN website and database: a new information system dedicated to the knottin scaffold. *Nucleic Acids Res.* 2004; 32:D156–159. [PubMed: 14681383]

48. Gracy J, Le-Nguyen D, Gelly JC, Kaas Q, Heitz A, Chiche L. KNOTTIN: the knottin or inhibitor cystine knot scaffold in 2007. *Nucleic Acids Res.* 2008; 36:D314–319. [PubMed: 18025039]

49. Wang CK, Kaas Q, Chiche L, Craik DJ. CyBase: a database of cyclic protein sequences and structures, with applications in protein discovery and engineering. *Nucleic Acids Res.* 2008; 36:D206–210. [PubMed: 17986451]

50. Daly NL, Rosengren KJ, Henriques ST, Craik DJ. NMR and protein structure in drug design: application to cyclotides and conotoxins. *Eur Biophys J.* 2011; 40:359–370. [PubMed: 21290122]

51. Wang CK, Hu SH, Martin JL, Sjogren T, Hajdu J, Bohlin L, et al. Combined X-ray and NMR analysis of the stability of the cyclotide cystine knot fold that underpins its insecticidal activity and potential use as a drug scaffold. *J Biol Chem.* 2009; 284:10672–10683. [PubMed: 19211551]

52. Svangard E, Goransson U, Smith D, Verma C, Backlund A, Bohlin L, et al. Primary and 3-D modelled structures of two cyclotides from *Viola odorata*. *Phytochemistry.* 2003; 64:135–142. [PubMed: 12946412]

53. Sze SK, Wang W, Meng W, Yuan R, Guo T, Zhu Y, et al. Elucidating the structure of cyclotides by partial acid hydrolysis and LC-MS/MS analysis. *Anal Chem.* 2009; 81:1079–1088. [PubMed: 19178340]

54. Craik DJ, Daly NL, Bond T, Waine C. Plant cyclotides: A unique family of cyclic and knotted proteins that defines the cyclic cystine knot structural motif. *J Mol Biol.* 1999; 294:1327–1336. [PubMed: 10600388]

55. Rosengren KJ, Daly NL, Plan MR, Waine C, Craik DJ. Twists, knots, and rings in proteins. Structural definition of the cyclotide framework. *J Biol Chem.* 2003; 278:8606–8616. [PubMed: 12482868]

56. Goransson U, Craik DJ. Disulfide mapping of the cyclotide kalata B1. Chemical proof of the cystic cystine knot motif. *J Biol Chem.* 2003; 278:48188–48196. [PubMed: 12960160]

57. Jagadish K, Camarero JA. Cyclotides, a promising molecular scaffold for peptide-based therapeutics. *Biopolymers.* 2010; 94:611–616. [PubMed: 20564025]

58. Daly NL, Gustafson KR, Craik DJ. The role of the cyclic peptide backbone in the anti-HIV activity of the cyclotide kalata B1. *FEBS Lett.* 2004; 574:69–72. [PubMed: 15358541]

59. Wang CK, Colgrave ML, Ireland DC, Kaas Q, Craik DJ. Despite a conserved cystine knot motif, different cyclotides have different membrane binding modes. *Biophys J.* 2009; 97:1471–1481. [PubMed: 19720036]

60. Plan MR, Saska I, Caguan AG, Craik DJ. Backbone cyclised peptides from plants show molluscicidal activity against the rice pest *Pomacea canaliculata* (golden apple snail). *J Agric Food Chem.* 2008; 56:5237–5241. [PubMed: 18557620]

61. Daly NL, Clark RJ, Plan MR, Craik DJ. Kalata B8, a novel antiviral circular protein, exhibits conformational flexibility in the cystine knot motif. *Biochem J.* 2006; 393:619–626. [PubMed: 16207177]

62. Hernandez JF, Gagnon J, Chiche L, Nguyen TM, Andrieu JP, Heitz A, et al. Squash trypsin inhibitors from *Momordica cochinchinensis* exhibit an atypical macrocyclic structure. *Biochemistry.* 2000; 39:5722–5730. [PubMed: 10801322]

63. Heitz A, Hernandez JF, Gagnon J, Hong TT, Pham TT, Nguyen TM, et al. Solution structure of the squash trypsin inhibitor MCOTI-II. A new family for cyclic knottins. *Biochemistry.* 2001; 40:7973–7983. [PubMed: 11434766]

64. Felizmenio-Quimio ME, Daly NL, Craik DJ. Circular proteins in plants: solution structure of a novel macrocyclic trypsin inhibitor from *Momordica cochinchinensis*. *J Biol Chem.* 2001; 276:22875–22882. [PubMed: 11292835]

65. Mylne JS, Colgrave ML, Daly NL, Chanson AH, Elliott AG, McCallum EJ, et al. Albumins and their processing machinery are hijacked for cyclic peptides in sunflower. *Nat Chem Biol.* 2011; 7:257–259. [PubMed: 21423169]

66. Saska I, Gillon AD, Hatsugai N, Dietzgen RG, Hara-Nishimura I, Anderson MA, et al. An asparaginyl endopeptidase mediates *in vivo* protein backbone cyclization. *J Biol Chem.* 2007; 282:29721–29728. [PubMed: 17698845]

67. Gillon AD, Saska I, Jennings CV, Guarino RF, Craik DJ, Anderson MA. Biosynthesis of circular proteins in plants. *Plant J.* 2008; 53:505–515. [PubMed: 18086282]

68. Ireland DC, Colgrave ML, Nguyencong P, Daly NL, Craik DJ. Discovery and characterization of a linear cyclotide from *Viola odorata*: implications for the processing of circular proteins. *J Mol Biol.* 2006; 357:1522–1535. [PubMed: 16488428]

69. Gruber CW, Cemazar M, Heras B, Martin JL, Craik DJ. Protein disulfide isomerase: the structure of oxidative folding. *Trends Biochem Sci.* 2006; 31:455–464. [PubMed: 16815710]

70. Marglin A, Merrifield RB. Chemical synthesis of peptides and proteins. *Annu Rev Biochem.* 1970; 39:841–866. [PubMed: 4920831]

71. Dawson PE, Muir TW, Clark-Lewis I, Kent SBH. Synthesis of Proteins by Native Chemical Ligation. *Science.* 1994; 266:776–779. [PubMed: 7973629]

72. Camarero JA, Pavel J, Muir TW. Chemical Synthesis of a Circular Protein Domain: Evidence for Folding-Assisted Cyclization. *Angew Chem Int Ed.* 1998; 37:347–349.

73. Tam JP, Lu Y-A. Synthesis of Large Cyclic Cystine-Knot Peptide by Orthogonal Coupling Strategy Using Unprotected Peptide Precursor. *Tetrahedron Lett.* 1997; 38:5599–5602.

74. Tam JP, Lu YA. A biomimetic strategy in the synthesis and fragmentation of a cyclic protein. *Prot Sci.* 1998; 7:1583–1592.

75. Gunasekera S, Daly NL, Clark RJ, Craik DJ. Dissecting the Oxidative Folding of Circular Cystine Knot Miniproteins. *Antioxid Redox Sign*. 2008; 11:971–980.10.1089/ars.2008.2295
76. Aboye TL, Clark RJ, Craik DJ, Göransson U. Ultra-stable peptide scaffolds for protein engineering-synthesis and folding of the circular cystine knotted cyclotide cycloviolacin O2. *Chembiochem*. 2008; 9:103–113. [PubMed: 18058973]
77. Camarero JA, Muir TW. Chemoselective backbone cyclization of unprotected peptides. *Chem Comm*. 1997; 1997:1369–1370.
78. Shao Y, Lu WY, Kent SBH. A novel method to synthesize cyclic peptides. *Tetrahedron Lett*. 1998; 39:3911–3914.
79. Camarero JA, Muir TW. Biosynthesis of a Head-to-Tail Cyclized Protein with Improved Biological Activity. *J Am Chem Soc*. 1999; 121:5597–5598.
80. Camarero JA, Mitchell AR. Synthesis of proteins by native chemical ligation using Fmoc-based chemistry. *Protein Pept Lett*. 2005; 12:723–728. [PubMed: 16305540]
81. Camarero JA, Cotton GJ, Adeva A, Muir TW. Chemical ligation of unprotected peptides directly from a solid support. *J Pept Res*. 1998; 51:303–316. [PubMed: 9560006]
82. Beligere GS, Dawson PE. Conformationally assisted ligation using C-terminal thioester peptides. *J Am Chem Soc*. 1999; 121:6332–6333.
83. Camarero JA, Adeva A, Muir TW. 3-Thiopropionic acid as a highly versatile multidetectable thioester resin linker. *Lett Pept Sci*. 2000; 7:17–21.
84. Ingenito R, Bianchi E, Fattori D, Pessi A. Solid phase synthesis of peptide C-terminal thioesters by Fmoc/t-Bu chemistry Source. *J Am Chem Soc*. 1999; 121:11369–11374.
85. Shin Y, Winans KA, Backes BJ, Kent SBH, Ellman JA, Bertozzi CR. Fmoc-Based Synthesis of Peptide-^aThioesters: Application to the Total Chemical Synthesis of a Glycoprotein by Native Chemical Ligation. *J Am Chem Soc*. 1999; 121:11684–11689.
86. Camarero JA, Hackel BJ, de Yoreo JJ, Mitchell AR. Fmoc-based synthesis of peptide alpha-thioesters using an aryl hydrazine support. *J Org Chem*. 2004; 69:4145–4151. [PubMed: 15176841]
87. Blanco-Canosa JB, Dawson PE. An efficient Fmoc-SPPS approach for the generation of thioester peptide precursors for use in native chemical ligation. *Angew Chem Int Ed Engl*. 2008; 47:6851–6855. [PubMed: 18651678]
88. Daly NL, Clark RJ, Craik DJ. Disulfide folding pathways of cystine knot proteins. Tying the knot within the circular backbone of the cyclotides. *J Biol Chem*. 2003; 278:6314–6322. [PubMed: 12482862]
89. Daly NL, Clark RJ, Göransson U, Craik DJ. Diversity in the disulfide folding pathways of cystine knot peptides. *Lett Pept Sci*. 2003; 10:523–531.
90. Cemazar M, Daly NL, Haggblad S, Lo KP, Yulyaningsih E, Craik DJ. Knots in rings. The circular knotted protein *Momordica cochinchinensis* trypsin inhibitor-II folds via a stable two-disulfide intermediate. *J Biol Chem*. 2006; 281:8224–8232. [PubMed: 16547012]
91. Aboye TL, Clark RJ, Burman R, Roig MB, Craik DJ, Göransson U. Interlocking Disulfides in Circular Proteins: Toward Efficient Oxidative Folding of Cyclotides. *Antioxid Redox Sign*. 2011; 14:77–86.
92. Marx UC, Korsinczky ML, Schirra HJ, Jones A, Condie B, Otvos L Jr, et al. Enzymatic cyclization of a potent bowman-birk protease inhibitor, sunflower trypsin inhibitor-1, and solution structure of an acyclic precursor peptide. *J Biol Chem*. 2003; 278:21782–21789. [PubMed: 12621047]
93. Perler FB, Adam E. Protein splicing and its applications. *Curr Opin Biotechnol*. 2000:377–383. [PubMed: 10975457]
94. Hirel PH, Schmitter MJ, Dessen P, Fayat G, Blanquet S. Extent of N-terminal methionine excision from *Escherichia coli* proteins is governed by the side-chain length of the penultimate amino acid. *Proc Natl Acad Sci U S A*. 1989; 86:8247–8251. [PubMed: 2682640]
95. Dwyer MA, Lu W, Dwyer JJ, Kossiakoff AA. Biosynthetic phage display: a novel protein engineering tool combining chemical and genetic diversity. *Chem Biol*. 2000; 7:263–274. [PubMed: 10780926]
96. Iwai H, Pluckthum A. Circular b-lactamase: stability enhancement by cyclizing the backbone. *FEBS Lett*. 1999:166–172. [PubMed: 10518012]

97. Cotton GJ, Ayers B, Xu R, Muir TW. Insertion of a Synthetic Peptide into a Recombinant Protein Framework; A Protein Biosensor. *J Am Chem Soc.* 1999; 121:1100–1101.

98. Camarero JA, Fushman D, Cowburn D, Muir TW. Peptide chemical ligation inside living cells: in vivo generation of a circular protein domain. *Bioorg Med Chem.* 2001; 9:2479–2484. [PubMed: 11553489]

99. Erlandson DA, Chytil M, Verdine GL. The leucine zipper domain controls the orientation of AP-1 in the NFAT AP-1 DNA complex. *Chem Biol.* 1996; 3:981–991. [PubMed: 9000009]

100. Baker RT, Smith SA, Marano R, McKee J, Board PG. Protein expression using cotranslational fusion and cleavage of ubiquitin. Mutagenesis of the glutathione-binding site of human Pi class glutathione S-transferase. *J Biol Chem.* 1994; 269:25381–25386. [PubMed: 7929235]

101. Varshavsky A. Ubiquitin fusion technique and related methods. *Methods Enzymol.* 2005; 399:777–799. [PubMed: 16338395]

102. Tolbert TJ, Wong C-H. New methods for proteomic research: preparation of proteins with N-terminal cysteines for labeling and conjugation. *Angew Chem Int Ed.* 2002; 41:2171–2174.

103. Liu D, Xu R, Dutta K, Cowburn D. N-terminal cysteiny1 proteins can be prepared using thrombin cleavage. *FEBS Lett.* 2008; 582:1163–1167. [PubMed: 18331839]

104. Dalbey RE, Lively MO, Bron S, van Dijl JM. The chemistry and enzymology of the type I signal peptidases. *Protein Sci.* 1997; 6:1129–1138. [PubMed: 9194173]

105. Paetzel M, Dalbey RE, Strynadka NC. Crystal structure of a bacterial signal peptidase apoenzyme: implications for signal peptide binding and the Ser-Lys dyad mechanism. *J Biol Chem.* 2002; 277:9512–9519. [PubMed: 11741964]

106. Hauser PS, Ryan RO. Expressed protein ligation using an N-terminal cysteine containing fragment generated in vivo from a pelB fusion protein. *Protein Expr Purif.* 2007; 54:227–233. [PubMed: 17493830]

107. Evans TC, Benner J, Xu M-Q. The *in Vitro* Ligation of Bacterially Expressed Proteins Using an Intein from *Metanobacterium thermoautotrophicum*. *J Biol Chem.* 1999; 274:3923–3926. [PubMed: 9933578]

108. Southworth MW, Amaya K, Evans TC, Xu MQ, Perler FB. Purification of proteins fused to either the amino or carboxy terminus of the *Mycobacterium xenopi* gyrase A intein. *Biotechniques.* 1999; 27:110–114. 116, 118–120. [PubMed: 10407673]

109. Mathys S, Evans TC, Chute IC, Wu H, Chong S, Benner J, et al. Characterization of a self-splicing mini-intein and its conversion into autocatalytic N- and C-terminal cleavage elements: facile production of protein building blocks for protein ligation. *Gene.* 1999; 231:1–13. [PubMed: 10231563]

110. Austin J, Kimura RH, Woo YH, Camarero JA. In vivo biosynthesis of an Ala-scan library based on the cyclic peptide SFTI-1. *Amino Acids.* 2010; 38:1313–1322. [PubMed: 19685144]

111. Shuker SB, Hajduk PJ, Meadows RP, Fesik SW. Discovering high-affinity ligands for proteins: SAR by NMR. *Science.* 1996; 274:1531–1534. [PubMed: 8929414]

112. Mylne JS, Craik DJ. 15N cyclotides by whole plant labeling. *Biopolymers.* 2008; 90:575–580. [PubMed: 18464271]

113. Colgrave ML, Kotze AC, Huang YH, O’Grady J, Simonsen SM, Craik DJ. Cyclotides: natural, circular plant peptides that possess significant activity against gastrointestinal nematode parasites of sheep. *Biochemistry.* 2008; 47:5581–5589. [PubMed: 18426225]

114. Colgrave ML, Huang YH, Craik DJ, Kotze AC. Cyclotide interactions with the nematode external surface. *Antimicrob Agents Chemother.* 2010; 54:2160–2166. [PubMed: 20211894]

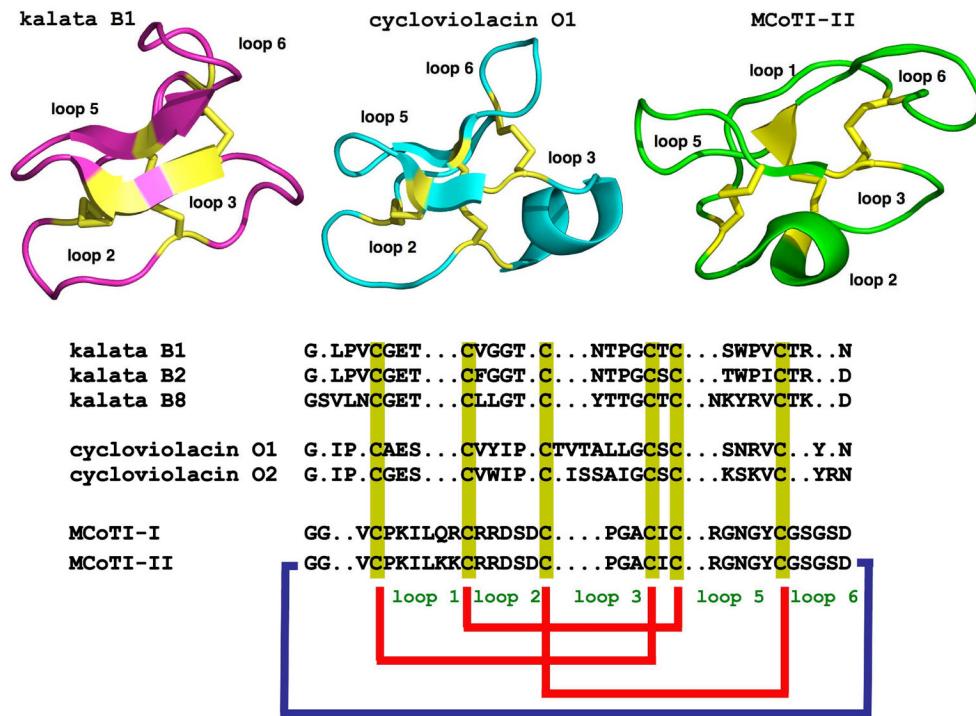
115. Barbetta BL, Marshall AT, Gillon AD, Craik DJ, Anderson MA. Plant cyclotides disrupt epithelial cells in the midgut of lepidopteran larvae. *Proc Natl Acad Sci U S A.* 2008; 105:1221–1225. [PubMed: 18202177]

116. Pranting M, Loov C, Burman R, Goransson U, Andersson DI. The cyclotide cycloviolacin O2 from *Viola odorata* has potent bactericidal activity against Gram-negative bacteria. *J Antimicrob Chemother.* 2010; 65:1964–1971. [PubMed: 20558471]

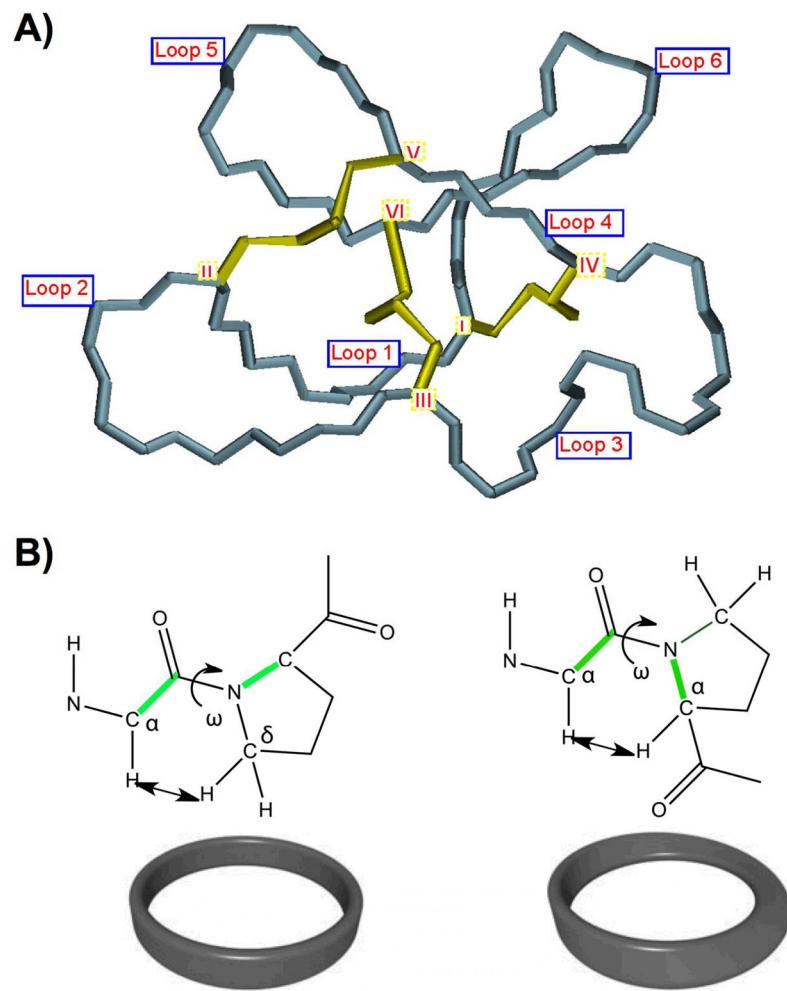
117. Henriques ST, Craik DJ. Cyclotides as templates in drug design. *Drug Discov Today.* 2010; 15:57–64. [PubMed: 19878736]

118. Chen B, Colgrave ML, Daly NL, Rosengren KJ, Gustafson KR, Craik DJ. Isolation and characterization of novel cyclotides from *Viola hederaceae*: solution structure and anti-HIV activity of vhl-1, a leaf-specific expressed cyclotide. *J Biol Chem.* 2005; 280:22395–22405. [PubMed: 15824119]
119. Wang CK, Colgrave ML, Gustafson KR, Ireland DC, Goransson U, Craik DJ. Anti-HIV cyclotides from the Chinese medicinal herb *Viola yedoensis*. *J Nat Prod.* 2008; 71:47–52. [PubMed: 18081258]
120. Ireland DC, Wang CK, Wilson JA, Gustafson KR, Craik DJ. Cyclotides as natural anti-HIV agents. *Biopolymers.* 2008; 90:51–60. [PubMed: 18008336]
121. Kamimori H, Hall K, Craik DJ, Aguilar MI. Studies on the membrane interactions of the cyclotides kalata B1 and kalata B6 on model membrane systems by surface plasmon resonance. *Anal Biochem.* 2005; 337:149–153. [PubMed: 15649388]
122. Shenkarev ZO, Nadezhdin KD, Lyukmanova EN, Sobol VA, Skjeldal L, Arseniev AS. Divalent cation coordination and mode of membrane interaction in cyclotides: NMR spatial structure of ternary complex Kalata B7/Mn²⁺/DPC micelle. *J Inorg Biochem.* 2008; 102:1246–1256. [PubMed: 18295894]
123. Shenkarev ZO, Nadezhdin KD, Sobol VA, Sobol AG, Skjeldal L, Arseniev AS. Conformation and mode of membrane interaction in cyclotides. Spatial structure of kalata B1 bound to a dodecylphosphocholine micelle. *FEBS J.* 2006; 273:2658–2672. [PubMed: 16817894]
124. Huang YH, Colgrave ML, Daly NL, Keleshian A, Martinac B, Craik DJ. The biological activity of the prototypic cyclotide kalata b1 is modulated by the formation of multimeric pores. *J Biol Chem.* 2009; 284:20699–20707. [PubMed: 19491108]
125. Svangard E, Goransson U, Hocaoglu Z, Gullbo J, Larsson R, Claeson P, et al. Cytotoxic cyclotides from *Viola tricolor*. *J Nat Prod.* 2004; 67:144–147. [PubMed: 14987049]
126. Herrmann A, Burman R, Mylne JS, Karlsson G, Gullbo J, Craik DJ, et al. The alpine violet, *Viola biflora*, is a rich source of cyclotides with potent cytotoxicity. *Phytochemistry.* 2008; 69:939–952. [PubMed: 18191970]
127. Svangard E, Burman R, Gunasekera S, Lovborg H, Gullbo J, Goransson U. Mechanism of action of cytotoxic cyclotides: cycloviolacin O2 disrupts lipid membranes. *J Nat Prod.* 2007; 70:643–647. [PubMed: 17378610]
128. Hoskin DW, Ramamoorthy A. Studies on anticancer activities of antimicrobial peptides. *Biochim Biophys Acta.* 2008; 1778:357–375. [PubMed: 18078805]
129. Herrmann A, Svangard E, Claeson P, Gullbo J, Bohlin L, Goransson U. Key role of glutamic acid for the cytotoxic activity of the cyclotide cycloviolacin O2. *Cell Mol Life Sci.* 2006; 63:235–245. [PubMed: 16389447]
130. Burman R, Herrmann A, Tran R, Kivela JE, Lomize A, Gullbo J, et al. Cytotoxic potency of small macrocyclic knot proteins: Structure-activity and mechanistic studies of native and chemically modified cyclotides. *Org Biomol Chem.* 2011; 9:4306–4314. [PubMed: 21491023]
131. Chen B, Colgrave ML, Wang C, Craik DJ. Cycloviolacin H4, a hydrophobic cyclotide from *Viola hederaceae*. *J Nat Prod.* 2006; 69:23–28. [PubMed: 16441062]
132. Simonsen SM, Sando L, Rosengren KJ, Wang CK, Colgrave ML, Daly NL, et al. Alanine scanning mutagenesis of the prototypic cyclotide reveals a cluster of residues essential for bioactivity. *J Biol Chem.* 2008; 283:9805–9813. [PubMed: 18258598]
133. Cascales L, Henriques ST, Kerr MC, Huang YH, Sweet MJ, Daly NL, et al. Identification and characterization of a new family of cell penetrating peptides: Cyclic cell penetrating peptides. *J Biol Chem.* 2011
134. Thongyoo P, Bonomelli C, Leatherbarrow RJ, Tate EW. Potent inhibitors of beta-tryptase and human leukocyte elastase based on the MCoTI-II scaffold. *Journal of medicinal chemistry.* 2009; 52:6197–6200. [PubMed: 19772295]
135. Sommerhoff CP, Avrutina O, Schmoldt HU, Gabrijelcic-Geiger D, Diederichsen U, Kolmar H. Engineered cystine knot miniproteins as potent inhibitors of human mast cell tryptase beta. *J Mol Biol.* 2010; 395:167–175. [PubMed: 19852971]
136. Gialeli C, Theocharis AD, Karamanos NK. Roles of matrix metalloproteinases in cancer progression and their pharmacological targeting. *FEBS J.* 2011; 278:16–27. [PubMed: 21087457]

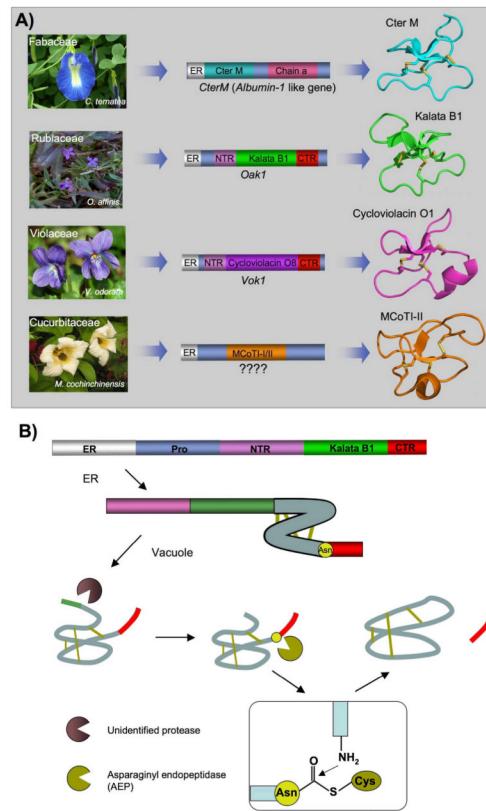
137. Giepmans BN, Adams SR, Ellisman MH, Tsien RY. The fluorescent toolbox for assessing protein location and function. *Science*. 2006; 312:217–224. [PubMed: 16614209]
138. You X, Nguyen AW, Jabaiah A, Sheff MA, Thorn KS, Daugherty PS. Intracellular protein interaction mapping with FRET hybrids. *Proc Natl Acad Sci U S A*. 2006; 103:18458–18463. [PubMed: 17130455]
139. Tam JP, Yu Q, Miao Z. Orthogonal ligation strategies for peptide and protein. *Biopolymers*. 1999; 51:311–332. [PubMed: 10685043]

**Figure 1.**

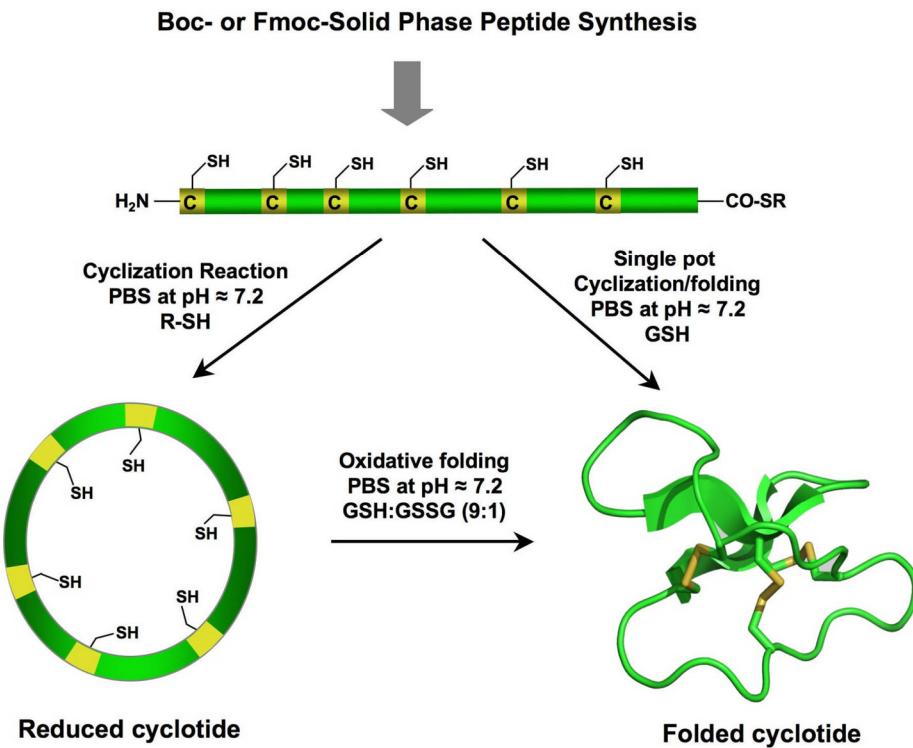
Primary and tertiary structures of cyclotides from the Möbius (kalata B1, pdb code: 1NB1), bracelet (cycloviolacin O1, pdb code: 1NBJ) and trypsin inhibitor (MCoTI-II, pdb code: 1IB9) subfamilies. The sequence of kalata B8, a novel hybrid cyclotide isolated from the plant *O. affinis* is also shown. Conserved cysteine residues are marked in yellow and disulfide connectivities in red. The circular backbone topology is shown with a blue line. Figure adapted from Ref. [4].

**Figure 2.**

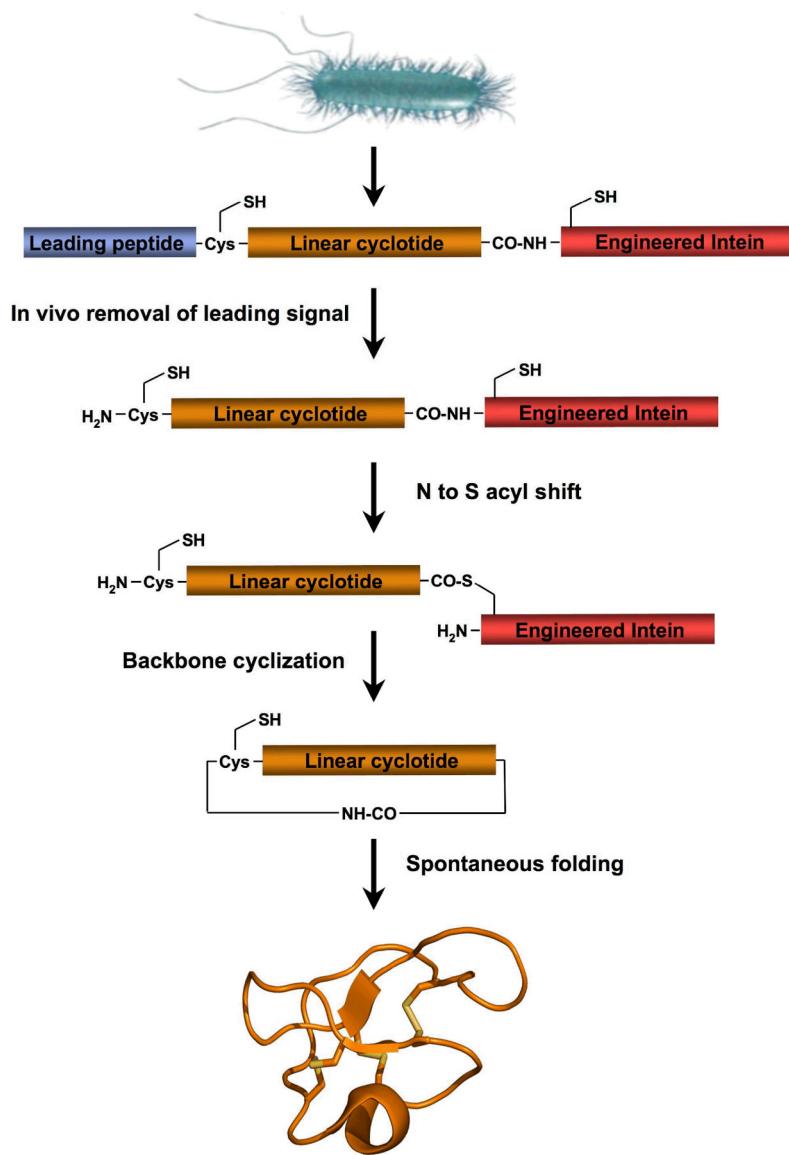
General features of the cyclic cystine knot (CCK) topology found in cyclotides. **A.** Detailed structure the cystine knot core and the connecting loops. The six Cys residues are labeled I through VI whereas loops connecting the different Cys residues are designated as loop 1 through 6, in numerical order from the N- to the C-terminus. **B.** Möbius (right) and bracelet (left) cyclotides are defined by the presence (Möbius) or absence (bracelet) of a Pro residue in loop 5 that introduces a twist in the circular backbone topology.

**Figure 3.**

Genetic origin and biosynthesis of cyclotides in plants. **A.** Rubiaceae and Violaceae plants have dedicated genes for the production of cyclotides [13]. These cyclotide precursors comprise an ER signal peptide, an N-terminal Pro region, the N-terminal repeat (NTR), the mature cyclotide domain and a C-terminal flanking region (CTR). Cyclotides from the Fabaceae family of plants isolated recently from *C. ternatea* [9, 10], show an ER signal peptide immediately followed by the cyclotide domain, which is flanked at the C-terminus by a peptide linker and the albumin a-chain. In this case, the cyclotide domain replaces albumin-1 b-chain. The genetic origin of the trypsin inhibitor cyclotides MCoTI-I/II (found in the seeds of *M. cochinchinensis*) remains yet to be determined. **B.** Scheme representing the proposed mechanism of protease-catalyzed cyclotidization. It has been proposed that the cyclization step is mediated by an asparaginyl endopeptidase (AEP), a common Cys protease found in plants. The cyclization takes place at the same time as the cleavage of the C-terminal pro-peptide from the cyclotide precursor protein through a transpeptidation reaction [66]. The transpeptidation reaction involves an acyl-transfer step from the acyl-AEP intermediate to the N-terminal residue of the cyclotide domain [67]. Figure adapted from Refs. [4, 11].

**Figure 4.**

Chemical synthesis of cyclotides by means of an intramolecular native chemical ligation (NCL). This approach requires the chemical synthesis of a linear precursor bearing an N-terminal Cys residue and an α -thioester moiety at the C-terminus. The linear precursor can be cyclized first under reductive conditions and then folded using a redox buffer containing reduced and oxidized glutathione (GSH). Alternatively, the cyclization and folding can be efficiently accomplished in a single pot reaction when the cyclization is carried out in the presence of reduced GSH as the thiol cofactor [18, 20].

**Figure 5.**

Biosynthetic approach for the recombinant expression of cyclotides using *E. coli* expression systems. Cleavage of the leading signal either in vitro [18] or in vivo [19, 20] by appropriate proteases provides the N-terminal Cys residue required for the cyclization. The backbone cyclization of the linear precursor is then mediated by a modified protein splicing unit or intein. Once the linear precursor is cyclized, folding is spontaneous for kalata B1 and MCoTI-I/II cyclotides [18–20].

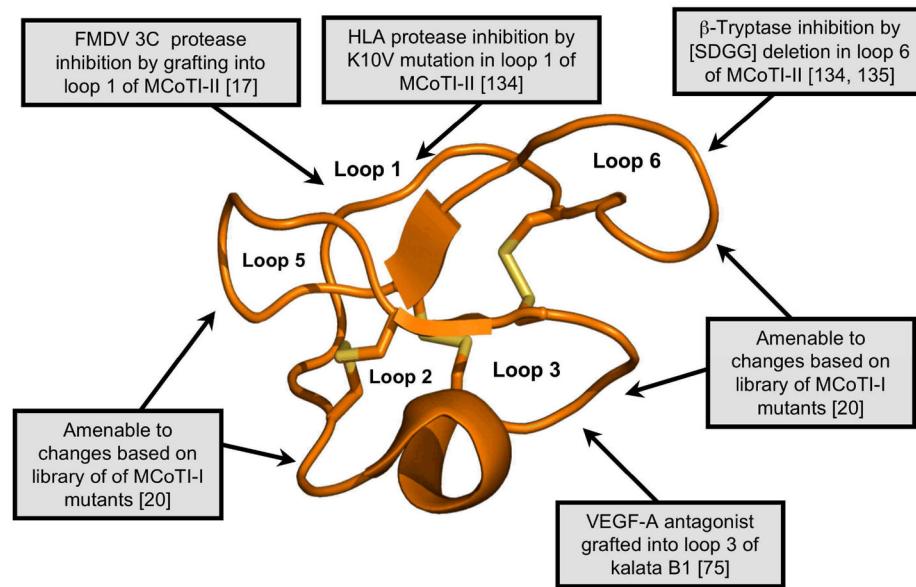


Figure 6.

Summary of some of the modifications made to kalata B1 and MCoTI-II to elicit novel biological activities. Modifications include the grafting of peptides onto various loops of kalata B1 [75] and MCoTI-II [17, 134, 135]. Cyclotide based libraries have been also generated using loops 1, 2, 3, 4 and 5 of MCoTI-I to study the effect of individual mutations on the biological activity and folding ability of MCoTI-I mutants [20]. The locations of the changes introduced into the cyclotide framework are illustrated using the MCoTI-II structure (pdb ID: 1IB9).

Biosynthesis and Antimicrobial Evaluation of Backbone-Cyclized α -Defensins

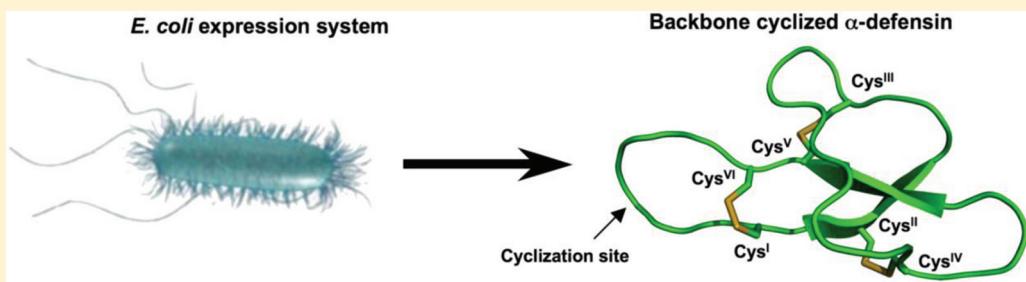
Angie E. Garcia,[†] Kenneth P. Tai,[§] Shadakshara S. Puttamadappa,^{||} Alexander Shekhtman,^{||} Andre J. Ouellette,[§] and Julio A. Camarero^{†,‡,*}

[†]Department of Pharmacology and Pharmaceutical Sciences, [‡]Department of Chemistry and

[§]Department of Pathology and Laboratory Medicine, University of Southern California, Los Angeles, California 90033, United States

^{||}Department of Chemistry, State University of New York, Albany, New York 12222, United States

 Supporting Information



ABSTRACT: Defensins are antimicrobial peptides that are important in the innate immune defense of mammals. Upon stimulation by bacterial antigens, enteric α -defensins are secreted into the intestinal lumen where they have potent microbicidal activities. Cryptdin-4 (Crp4) is an α -defensin expressed in Paneth cells of the mouse small intestine and the most bactericidal of the known cryptdin isoforms. The structure of Crp4 consists of a triple-stranded antiparallel β -sheet but lacks three amino acids between the fourth and fifth cysteine residues, making them distinct from other α -defensins. The structure also reveals that the α -amino and C-terminal carboxylic groups are in the proximity of each other ($d \approx 3 \text{ \AA}$) in the folded structure. We present here the biosynthesis of backbone-cyclized Crp4 using a modified protein splicing unit or intein. Our data show that cyclized Crp4 can be biosynthesized by using this approach both *in vitro* and *in vivo*, although the expression yield was significantly lower when the protein was produced inside the cell. The resulting cyclic defensins retained the native α -defensin fold and showed equivalent or better microbicidal activities against several Gram-positive and Gram-negative bacteria when compared to native Crp4. No detectable hemolytic activity against human red blood cells was observed for either native Crp4 or its cyclized variants. Moreover, both forms of Crp4 also showed high stability to degradation when incubated with human serum. Altogether, these results indicate the potential for backbone-cyclized defensins in the development of novel peptide-based antimicrobial compounds.

The cyclization of peptides has commonly been used to generate more active and stable scaffolds for therapeutic purposes.^{1–6} In general, backbone-cyclized peptides are more resistant than linear peptides to chemical, thermal, and enzymatic degradation.⁷ Because circular peptides have no N- and C-termini, they are more rigid, endowing them with the ability to persist in physiologic environments such as blood serum.^{3,6}

The production of backbone-cyclized or circular peptides can be performed chemically by using solid-phase peptide synthesis in combination with native chemical ligation^{8–13} or recombinantly in bacteria by using modified protein splicing units or inteins.^{14–16} The latter method has been used for the biosynthesis of several disulfide-rich cyclic peptides such as cyclotides^{17,18} and sunflower trypsin inhibitor 1 (SFTI-1).¹⁹ The expression of cyclic peptides *in vivo* has many applications, including the generation of peptide libraries for high-throughput screening of biological activities such as antimicrobial activity or

specific inhibition of protein–protein interactions involved in various pathologies.

Mammalian defensins are a family of disulfide-stabilized, host defense peptides.^{20–24} They are classically known for their antimicrobial activities but play a role in many other defense mechanisms, including wound healing,^{25,26} immune modulation,^{27–30} neutralization of endotoxin,^{31–33} and anticancer activities.³⁴ Mammalian defensins are cationic peptides with largely β -sheet structures and six conserved cysteines, which form three intramolecular disulfide bonds (Figure 1). They are divided into three structurally distinct groups, α -, β -, and θ -defensins. The overall fold of α - and β -defensins is quite similar despite differences in disulfide connectivities, and the presence of an N-terminal α -helix segment in β -defensins that is lacking

Received: September 14, 2011

Revised: October 27, 2011

Published: October 31, 2011

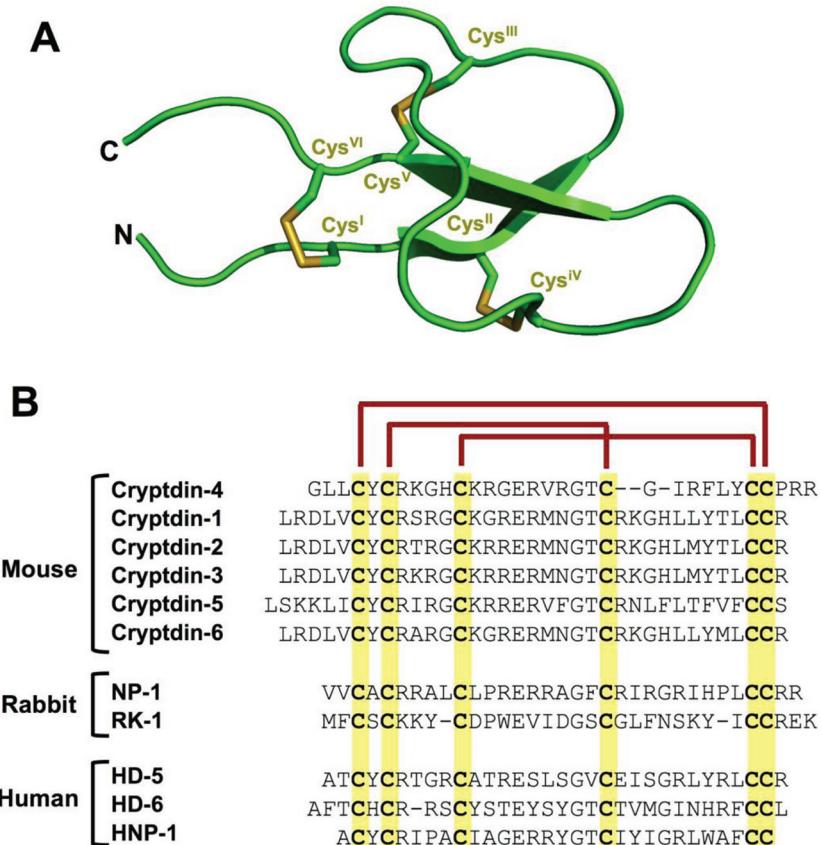


Figure 1. (A) Tertiary structure of Crp4 (Protein Data Bank entry 1TV0).³⁵ (B) Primary structures of representative α -defensins from mouse, rabbit, and human. The six conserved cysteines are highlighted, and the canonical 1–6, 2–4, and 3–5 disulfide connectivities are shown at the top of the sequences.

in α -defensins.³⁵ θ -Defensins, on the other hand, are cyclic peptides formed by the head-to-tail covalent assembly of two nonapeptides and, to date, are the only known cyclic polypeptides expressed in animals.²²

The antibacterial activity of α - and β -defensins is highly dependent on the ionic strength of the media, and salt-dependent inactivation of defensins in patients with cystic fibrosis has been proposed as the potential cause of chronic pulmonary infections in these patients.³⁶ The antimicrobial activity of the naturally cyclic θ -defensins, in contrast, has been shown to be less sensitive to salt concentration than those of α - and β -defensins.²² This difference has been attributed to the circular structure of θ -defensins because the acyclic forms are more salt sensitive.²² In agreement with this, replacement of Cys pairs with Gly in the rabbit α -defensin neutrophil peptide 1 (NP-1) and backbone cyclization provided biologically active defensin analogues that were less sensitive to salt.³⁷ However, no studies of the stability of these analogues were performed.

Intrigued by these results, we decided to investigate further the effects of backbone cyclization on the activity and stability of fully folded α -defensins. In this work, we have used the α -defensin cryptdin-4 (Crp4) from murine Paneth cells as a model system.³⁸ Cryptdins are secreted into the crypt lumen by mouse small intestinal Paneth cells^{39–42} where they exert potent microbicidal effects and determine the composition of the ileal microbiome.^{38,43–45} Cryptdins exhibit very similar antibacterial activity against Gram-positive and Gram-negative bacteria, with Crp4 displaying the greatest microbicidal activity in *in vitro* assays.⁴⁶ Cryptdins are classified as enteric defensins,

and homologues have been found in other mammals.^{42,47} The structure of Crp4 consists of a triple-stranded antiparallel β -sheet³⁵ but lacks three amino acids between the fourth and fifth cysteine residues, making them distinct from other α -defensins (Figure 1). The structure also reveals that the α -amino and C-terminal carboxylic groups are proximal in the folded structure (Figure 1), suggesting that the introduction of backbone cyclization at this point could help to stabilize the defensin scaffold without disrupting its structure and biological activity.

Here we present the recombinant expression and characterization of several backbone cyclized analogues of fully folded Crp4. Backbone cyclization was performed *in vitro* and *in vivo* using an intramolecular intein-mediated cyclization in *Escherichia coli* cells.⁴⁸ The expression yields of folded cyclized Crp4 *in vivo* were, however, lower than those obtained when the cyclization and folding were performed *in vitro* from the intein fusion precursors. Our data also show that cyclized versions of Crp4 adopt a native folded structure and exhibit equal or better microbicidal activities against several Gram-positive and Gram-negative bacteria than native Crp4 (Crp4-wt). Native and cyclized variants of Crp4 were antimicrobial in the presence of human serum, and the biological activity of cyclized Crp4 was slightly less affected by the presence of 160 mM sodium chloride than native Crp4. Also, neither native nor cyclized Crp4 showed detectable hemolytic activity against human red blood cells. Moreover, both native and cyclized Crp4 showed remarkable resistance to degradation in 100% human serum with half-life values of >48 h. Altogether, these results indicate that backbone cyclization provides a means for engineering

defensins to improve biological activity while providing excellent stability to serum degradation, highlighting the potential of these peptide scaffolds for the development of novel antimicrobial compounds.

■ EXPERIMENTAL PROCEDURES

Analytical reverse-phase HPLC (RP-HPLC) was performed on an HP1100 series instrument with 220 and 280 nm detection using a Vydac C4 or C18 column (5 mm, 4.6 mm × 150 mm) at a flow rate of 1 mL/min. Semipreparative RP-HPLC was performed on a Waters Delta Prep system fitted with a Waters 2487 ultraviolet-visible (UV-vis) detector using a Vydac C4 column (5 μm, 10 mm × 250 mm) at a flow rate of 5 mL/min. All runs used linear gradients of 0.1% aqueous trifluoroacetic acid (TFA, solvent A) versus 0.1% TFA and 90% MeCN in H₂O (solvent B). UV-vis spectroscopy was conducted on an Agilent 8453 diode array spectrophotometer. Electrospray mass spectrometry (ES-MS) was performed on an Applied Biosystems API 3000 triple quadrupole mass spectrometer. Calculated masses were obtained by using ProMac version 1.5.3. Protein samples were analyzed on sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) 4 to 20% Tris-glycine gels (Lonza, Rockland, ME). The gels were then stained with Pierce (Rockford, IL) Gelcode Blue, photographed and digitized using a Kodak (Rochester, NY) EDAS 290 instrument, and quantified using NIH Image-J (<http://rsb.info.nih.gov/ij/>). DNA sequencing was performed by the DNA Sequencing and Genetic Analysis Core Facility at the University of Southern California using an ABI 3730 DNA sequencer, and the sequence data were analyzed with DNASTar (Madison, WI) Lasergene version 8.0.2. All chemicals were obtained from Sigma-Aldrich (Milwaukee, WI) unless otherwise indicated.

Cloning and in Vitro Expression of Backbone-Cyclized Crp4 Variants. Synthetic DNA oligos (Integrated DNA Technologies, Coralville, IA) encoding the different backbone-cyclized Crp4 analogues (Table S1 of the Supporting Information) were annealed and ligated into the pTXB1 vector (New England Biolabs, Ipswich, MA) using the NdeI and SapI restriction sites as described previously.^{17,49} The resulting plasmids were transformed into either BL21(DE3) or Origami2(DE3) cells (EMD Chemicals, Gibbstown, NJ) and grown in LB broth. Transformed BL21(DE3) cells were induced with 0.3 mM IPTG for 4 h at 30 °C and transformed Origami2(DE3) cells with 0.1 mM IPTG for 20 h at 22 °C. Cells were lysed in 0.1 mM EDTA, 1 mM PMSF, 50 mM sodium phosphate, and 250 mM sodium chloride buffer (pH 7.2) containing 5% glycerol by sonication. The soluble fraction was incubated with chitin beads (New England Biolabs) for 1 h at 4 °C, and the beads were washed with column buffer [0.1 mM EDTA, 50 mM sodium phosphate, and 250 mM sodium chloride buffer (pH 7.2)] containing 0.1% Triton X-100 followed by washes with column buffer without Triton X-100. The peptide was cyclized and folded in vitro using column buffer at pH 7.2 containing 50–100 mM reduced glutathione (GSH) for 2–3 days at room temperature with gentle rocking. We found that under these conditions backbone-cyclized Crp4 variants bind strongly to the chitin column and therefore were eluted using 8 M GdmCl in water. The corresponding supernatant and washes were pooled, and the backbone-cyclized Crp4 peptides were purified by semipreparative HPLC using a linear gradient from 20 to 40% solvent B over 30 min. Purified products were characterized by HPLC and ES-MS. All Crp4 variants were quantified by UV-vis using a molar

absorption coefficient of 3365 M⁻¹ cm⁻¹. The expression of Crp4–intein fusion precursors was quantified by first desorption of the proteins from an aliquot of chitin beads using 8 M GdmCl and then measurement by UV-vis using a molar absorption coefficient of 39015 M⁻¹ cm⁻¹.

Preparation of Crp4-R/A. The linear reduced and alkylated Crp4 (Crp4-R/A) was produced by expression of the Crp4-1 intein precursor in BL21(DE3) cells using the induction conditions described above. After purification of the fusion with chitin beads, Crp4-1 was cleaved from the intein using 100 mM NH₂OH in water at pH 7.2 for 18 h at room temperature. The resulting linear Crp4-1 peptide was reduced with 5 mM DTT at 37 °C for 3 h and alkylated with 12.5 mM iodoacetamide for 10 min at room temperature. Crp4-R/A was purified by semipreparative HPLC as described above. The purified product was characterized by HPLC and ES-MS (Figure S7 of the Supporting Information).

Expression of Native Crp4-wt and Crp4-6C/A. Native Crp4-wt and Crp4-6C/A were expressed in *E. coli*, purified, and refolded as previously described.^{50,51}

Expression of ¹⁵N-Labeled Backbone-Cyclized Crp4 Variants. Expression was conducted using BL21(DE3) cells as described above except they were grown in M9 minimal medium containing 0.1% ¹⁵NH₄Cl as the nitrogen source.^{18,52} Cyclization and folding were performed as described above. The ¹⁵N-labeled backbone-cyclized Crp4 defensins were purified by semipreparative HPLC as described above. Purified products were characterized by HPLC and ES-MS (Figure S3 of the Supporting Information).

In Vivo Expression of Crp4-1. Origami2(DE3) cells transformed with the plasmid encoding the intein precursor of Crp4-1 were induced, harvested, and lysed as described above. The insoluble pellet was first washed three times with column buffer containing 0.1% Triton X-100 and then twice with just column buffer. The resulting pellet was dissolved in a minimal amount of 8 M GdmCl in water. Both the soluble cell lysate and the solubilized cell lysate pellet were extracted using C18 SepPak cartridges (Waters, Milford, MA) with elution in a MeCN/H₂O mixture (3:2, v/v) containing 0.1% TFA. The samples were analyzed by HPLC and tandem mass spectrometry (HPLC-MS/MS) using a C18-HPLC column (5 mm, 2.1 mm × 100 mm), and H₂O/MeCN buffers containing 0.1% formic acid as the mobile phase. Typical analysis used a linear gradient from 0 to 90% MeCN in H₂O over 10 min. Detection was performed on an API 3000 ES-MS instrument using a multiple-reaction monitoring (MRM) mode. Data were collected and processed using Analyst (Applied Biosystems). The calibration curve using purified cyclic Crp4-1 was found to be linear in the range of 5–50 ng.

NMR Spectroscopy. NMR samples were prepared by dissolving ¹⁵N-labeled backbone-cyclized Crp4 variants in 80 mM potassium phosphate in a 90% H₂O/10% ²H₂O mixture (v/v) or 100% D₂O to a concentration of approximately 0.2 mM with the pH adjusted to 4.5 or 6.0 by addition of dilute HCl. All ¹H NMR data were recorded on a Bruker Avance II 700 MHz spectrometer equipped with a cryoprobe. Data were acquired at 27 °C, and 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) was used as an internal reference. All three-dimensional (3D) experiments, ¹H{¹⁵N} TOCSY-HSQC and ¹H{¹⁵N} NOESY, were performed according to standard procedures⁵³ with spectral widths of 12 ppm in the proton dimension and 35 ppm in the nitrogen dimension. The carrier frequency was centered on the water signal, and the solvent was suppressed by

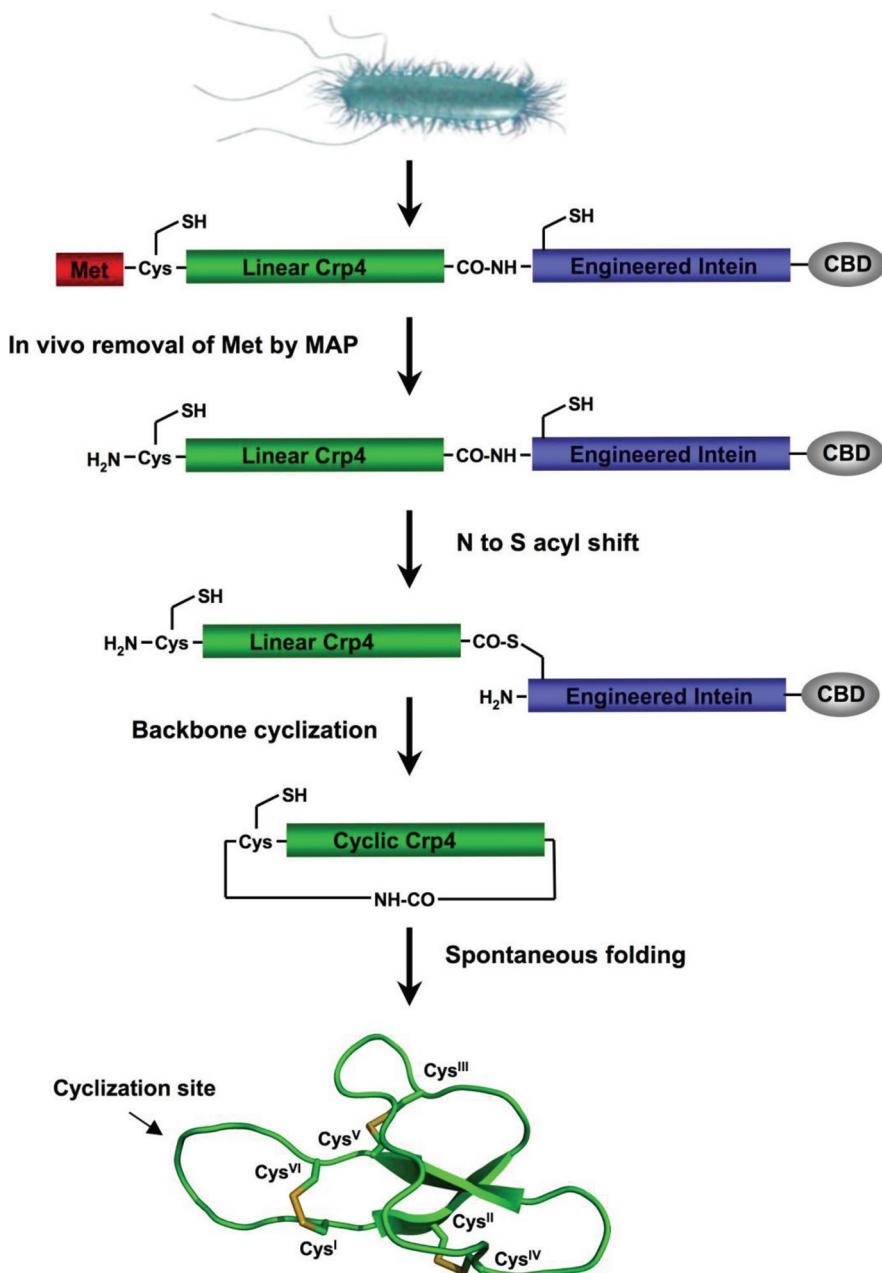


Figure 2. Biosynthesis of backbone-cyclized Crp4 variants using a modified intein in combination with native chemical ligation. CBD stands for chitin binding domain.

using a WATERGATE pulse sequence. TOCSY (spin-lock time of 80 ms) and NOESY (mixing time of 150 ms) spectra were recorded using 1024 t_3 points, 256 t_2 blocks, and 128 t_1 blocks of 16 transients. Spectra were processed using Topspin 1.3 (Bruker). Each 3D data set was apodized by a 90°-shifted sine bell-squared process in all dimensions and zero-filled to 1024 \times 512 \times 256 points prior to Fourier transformation. Assignments for the backbone nitrogens, H^a and H' protons (Figures S4 and Tables S2 and S3 of the Supporting Information), were obtained using standard procedures.^{53,54}

Bactericidal Peptide Assays. Native Crp4 (Crp4-wt) and backbone-cyclized Crp4 variants were tested for bactericidal peptide activity against Gram-negative and Gram-positive bacteria. Bacteria growing exponentially in trypticase soy broth (TSB) were centrifuged at 10000g for 3 min and washed three times with 10 mM PIPES buffer (pH 7.4) supplemented

with 1% (v/v) TSB (10 mM PIPES-TSB). Approximately 1–5 \times 10⁶ colony-forming units (CFU) per milliliter of bacteria were incubated with peptides in a total volume of 50 μ L of 10 mM PIPES-TSB. In assays performed in the presence of salt or serum, sample mixtures were incubated in 10 mM PIPES-TSB supplemented with sodium chloride or heat-inactivated human serum, respectively. Mixtures of bacteria and peptide were incubated at 37 °C with shaking for 1 h, and 20 μ L aliquot samples were diluted in 2 mL of 10 mM PIPES buffer (pH 7.4) and plated on TSB plates using an Autoplate 4000 instrument (Spiral Biotech Inc., Bethesda, MD). After incubation overnight at 37 °C, bacterial cell survival was assessed by counting CFUs.

Hemolysis Assay. EDTA-anticoagulated human blood was obtained from a healthy donor (Bioreclamation, LLC, Hicksville, NY) and centrifuged at 234g for 10 min at 22 °C. Red blood cells (RBCs) were washed four times with Dulbecco's

phosphate-buffered saline (DPBS) (Mediatech Inc., Manassas, VA) containing 4 mM EDTA and resuspended in DPBS without EDTA. Peptides diluted in DPBS to the concentrations shown were assayed for hemolysis in triplicate by incubation with 2% (v/v) RBCs for 1 h at 37 °C in an atmosphere of 5% CO₂. The cells were centrifuged at 234g for 10 min at 22 °C, and the absorbencies of the supernatants were measured at 405 nm. The hemolytic activity of each peptide was calculated relative to the 100% hemolysis obtained by incubation of RBCs with 1% Triton X-100.

Serum Stability Assay. Peptides at 100 μ g/mL (~27 μ M) in human serum (Lonza) were incubated at 37 °C. After various time points, triplicate serum aliquots were removed, quenched with 6 M urea in water, and incubated for 10 min at 4 °C. Subsequently, serum proteins were precipitated with 20% trichloroacetic acid for 10 min at 4 °C and centrifuged at 13000 rpm for 10 min at 4 °C. The supernatants were analyzed by C4 RP-HPLC, and the pellets were dissolved in 8 M GdmCl and also analyzed by HPLC. The percentage of peptide recovery was determined by integration of the HPLC peaks at 220 nm. Peptide identity was also confirmed by ES-MS.

RESULTS

Design of Backbone-Cyclized Crp4 Defensins. The biosynthesis of backbone-cyclized Cys-rich polypeptides using modified protein-splicing units or inteins has previously been demonstrated for several cyclotides^{17,18,49,52} and sunflower trypsin inhibitor 1 (SFTI-1).⁵⁵ Here, we have used a similar strategy for the biosynthesis of several backbone-cyclized Crp4 variants in *E. coli* cells. This approach makes use of a modified intein in combination with an intramolecular native chemical ligation reaction (NCL) (Figure 2).^{17,49} Intramolecular NCL requires the presence of an N-terminal Cys residue and a C-terminal α -thioester group in the same linear precursor molecule.^{8,14} For this purpose, the corresponding Crp4 linear precursor was fused in frame at the N- and C-termini to a Met residue and a modified *Mxe* Gyrase A intein, respectively. This allows the generation of the required C-terminal thioester and N-terminal Cys residue after in vivo processing by endogenous Met aminopeptidase (MAP) (Figure 2). We designed five linear Crp4 precursors (Crp4-1–Crp4-5) to explore the best ligation site as well as the best linker required for optimal cyclization and folding of the resulting backbone-cyclized Crp4 variants (Scheme 1 and Table 1). We first decided to link

Scheme 1. Sequences of the Different Crp4–Intein Precursors Used in This Study^a

Intein precursor	Sequence
Crp4-1	CYCRKGHCKRGERVRGTCGIRFLYCCP ^{RR} <u>GLL</u> -intein
Crp4-2	CYCRKGHCKRGERVRGTCGIRFLYCCP ^{RR} <u>GLL</u> -intein
Crp4-3	CYCRKGHCKRGERVRGTCGIRFLYCCP ^{RR} <u>GPGLL</u> -intein
Crp4-4	CRKGHCKRGERVRGTCGIRFLYCCP ^{RR} <u>GLL</u> CY-intein
Crp4-5	CCP ^{RR} <u>GLL</u> CYCRKGHCKRGERVRGTCGIRFLY-intein

^aThe first three residues located at the N-terminus of Crp4-wt are underlined in blue for reference. The residues added to the ligation site to facilitate cyclization are colored green.

together through a peptide bond the native N- and C-terminal residues of Crp4-wt, Gly1 and Arg32, respectively. These two residues are in the proximity of each other in the folded structure of Crp4-wt, and in principle, the formation of this bond should not introduce too much conformational stress in

Table 1. Amino Acid Sequences of the Crp4 Peptides Used in This Study

Peptide	Sequence
Crp4-wt	<u>GLL</u> CYCRKGHCKRGERVRGTCGIRFLYCCP ^{RR} ^a
Crp4-6C/A	<u>GLL</u> A ^Y ARKGH <u>A</u> KRGERVRGTA <u>G</u> IRFLYAA ^{PPR} ^b
Crp4-1	cyclo[CYCRKGHCKRGERVRGTCGIRFLYCCP ^{RR} <u>GLL</u>] ^c
Crp4-2	cyclo[CYCRKGHCKRGERVRGTCGIRFLYCCP ^{RR} <u>GGLL</u>] ^c
Crp4-3	cyclo[CYCRKGHCKRGERVRGTCGIRFLYCCP ^{RR} <u>GPGLL</u>] ^c
Crp4-R/A	<u>CYCRKGHCKRGERVRGTCGIRFLYCCP^{RR}GLL</u> ^d

^aThe first three amino acids of the Crp4 natural N-terminus are colored blue for reference. ^bAla residues replacing original Cys residues are colored magenta. ^cThe linkers added to assist in cyclization are colored green. ^dCys residues reduced and alkylated are underlined.

the newly formed loop. To facilitate the ligation, we used Cys¹ (Crp4-1), Cys^{II} (Crp4-4), and Cys^V (Crp4-5) as the N-terminal residues in the corresponding linear precursors (Table 1). In addition, we also explored the effect of adding extra residues in the newly formed loop in the backbone-cyclized Crp4. Hence, we added an extra Gly residue (Crp4-2) and a Pro-Gly sequence (Crp4-3) between the native N- and C-terminal residues, Gly1 and Arg32, respectively (Table 1). These two additions should increase the flexibility of the new loop as well as facilitate the formation of the required turn.⁵⁶ In these two variants, we used only Cys¹ as the N-terminal residue to allow for the cyclization reaction.

In Vitro Biosynthesis of Backbone-Cyclized Crp4 Peptides. All five linear Crp4–intein fusion precursors were expressed in BL21(DE3) cells at 30 °C and purified by affinity chromatography using chitin–Sepharose beads. The Crp4–intein precursors have a chitin binding domain (CBD) fused at the C-terminus of the intein domain to facilitate purification. The expression yields of the different Crp4–intein constructs in the soluble cell lysate were estimated by UV spectroscopy and ranged from \approx 5 mg/L for intein precursor Crp4-1 to \approx 14 mg/L for precursor Crp4-5 (Table 2). The rest of the

Table 2. Expression Yields and Cleavage Percentages of Crp4–Intein Precursors in BL21(DE3) Cells and Yields of Cyclic and Folded Crp4 Variants after in Vitro Cleavage of the Intein

precursor	yield (mg/L)		% of soluble intein cleaved		yield ^c [μ g/L (%)] of cyclic/folded
	S ^a	P ^b	in vivo	in vitro	
Crp4-1	5.5	142.1	23.9	84.2	200 (50%) ^d
Crp4-2	9.8	24.7	22.9	73.4	120 (20%) ^d
Crp4-3	10.5	94.3	20.7	70.4	174 (27%) ^d
Crp4-4	9.0	0	29.1	67.7	43 (10%) ^d
Crp4-5	13.9	0	7.5	45.6	35 (5%) ^d

^aSoluble cell lysate fraction. ^bInsoluble cell lysate fraction. ^cObtained after in vitro cleavage with GSH (see Experimental Procedures).

^dCyclization/folding yield based on the amount of precursor cleaved during the GSH-induced in vitro cleavage.

constructs gave similar expression yields in the soluble fraction (\approx 10 mg/L). All of the linear precursors using Cys¹ as the N-terminal residue (Crp4-1, Crp4-2, and Crp4-3) also produced a

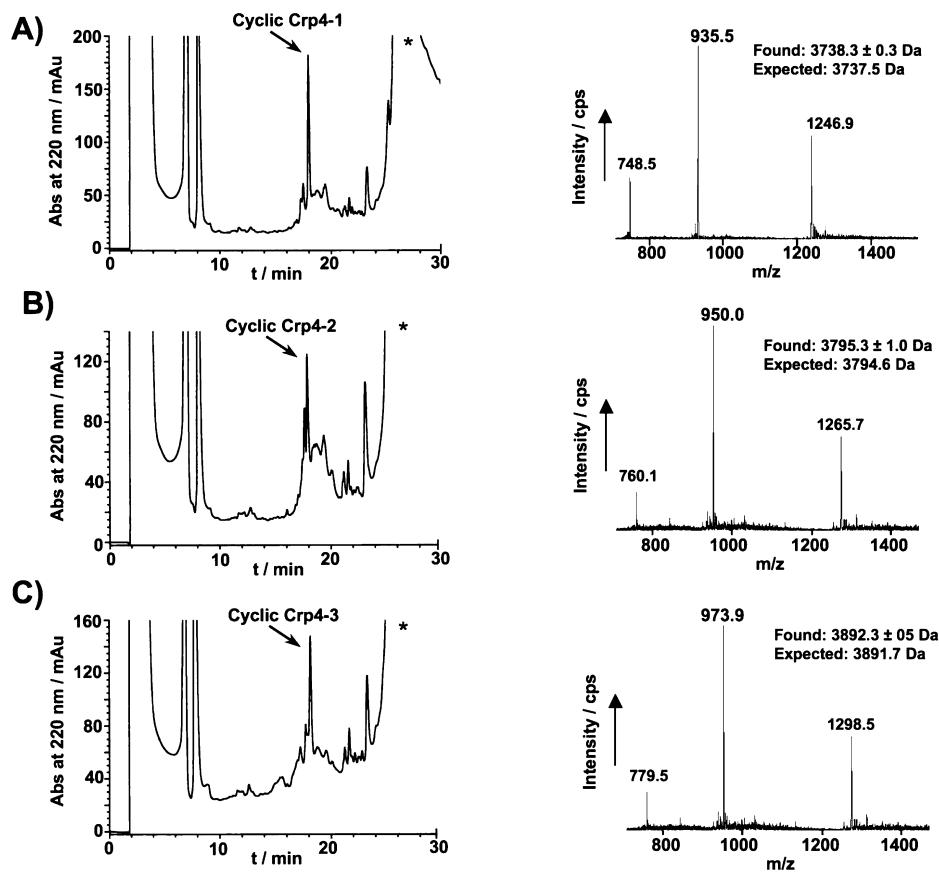


Figure 3. In vitro GSH-induced cyclization of precursors Crp4-1 (A), Crp4-2 (B), and Crp4-3 (C). On the left are the RP-HPLC chromatograms of the crude reaction and on the right the ES-MS data of the corresponding cyclic folded product indicated with an arrow. The large peak labeled with an asterisk is the intein–CBD protein. Experimental and theoretical molecular weights are shown with MS spectra.

significant amount of protein in the insoluble fraction ranging from 142 mg/L for Crp4-1 (\approx 90% of the total precursor expressed) to 24 mg/L for Crp4-2 (\approx 70% of the total precursor expressed). Interestingly, neither Crp4-4 nor Crp4-5 showed any significant expression in the insoluble fraction (Figure S1 of the Supporting Information), suggesting that the arrangement of the Crp4 peptide primary sequence plays an important role in the folding of the precursors and subsequent shuttling to inclusion bodies. Analysis of the purified fusion proteins by SDS–PAGE (Figure S1 of the Supporting Information) revealed that all of the linear precursors except Crp4-5 showed \approx 20% cleavage of the intein fusion in vivo when expressed in BL21(DE3) cells for 4 h at 30 °C. In contrast, precursor Crp4-5 gave only 7% in vivo cleavage.

We next tested the ability of the different precursors to be cleaved in vitro using reduced glutathione (GSH). GSH has been shown to promote cyclization and concomitant folding when used in the biosynthesis of Cys-rich cyclic polypeptides.^{18,49,55} The cyclization/folding reaction was performed on the chitin beads where the corresponding precursors had been purified. The best cleavage/cyclization conditions were achieved using 100 mM GSH in phosphate buffer at pH 7.2 for 48 h. Under these conditions, \approx 85% of precursor Crp4-1 was cleaved in vitro (Table 2 and Figure S1 of the Supporting Information). Precursors Crp4-2 and Crp4-3 were also cleaved efficiently (\approx 70%) under these conditions. In contrast, linear precursors Crp4-4 and Crp4-5 were cleaved less efficiently by GSH, \approx 65% for Crp4-4 and \approx 45% for Crp4-5. HPLC analysis of the crude cyclization mixture revealed in all the cases the

main peptide product was the corresponding folded backbone-cyclized Crp4 variant as revealed by ES-MS analysis (Figure 3 and Figure S2 of the Supporting Information). Other peptide peaks in the HPLC chromatograms were identified as incorrectly folded GSH adducts. Among the different linear precursors, Crp4-1 gave the best cyclization/folding yield (Figure 3A), producing \sim 200 μ g/L cyclized Crp4-1 [\approx 50% of the theoretical yield (see Table 1)]. The cyclization/folding crude mixture for Crp4-2 and Crp4-3 (Figure 3B,C), both using the same ligation site as Crp4-1 but with different linker lengths for the new loop formed at the ligation site, gave lower yields (\approx 120 and 170 μ g/L, respectively) than Crp4-1. The HPLC trace for the cyclization reaction of precursor Crp4-3 was, however, cleaner than that of Crp4-2, indicating the cyclization/folding was more efficient in that precursor (Figure 3B,C). The cyclization/folding yields for Crp4-4 and Crp4-5 (Figure S2 of the Supporting Information) were similar and estimated by HPLC to be \sim 40 and \sim 35 μ g/L, respectively (Table 2). These lower yields could be attributed, in part, to the less efficient cleavage of the intein in these precursors.

It is worth noting that during the cleavage of the intein precursors we found that the corresponding cyclized Crp4 peptides were able to bind strongly to the chitin beads under the conditions used for the GSH-induced cyclization/folding and required the use of 8 M GdmCl to elute the cyclized peptides from the solid support. This treatment also eluted the intein–CBD protein byproduct (Figure 3). Accordingly, we tested whether the cyclized Crp4 variants were binding to the chitin–Sepharose beads or intein–CBD protein by incubation

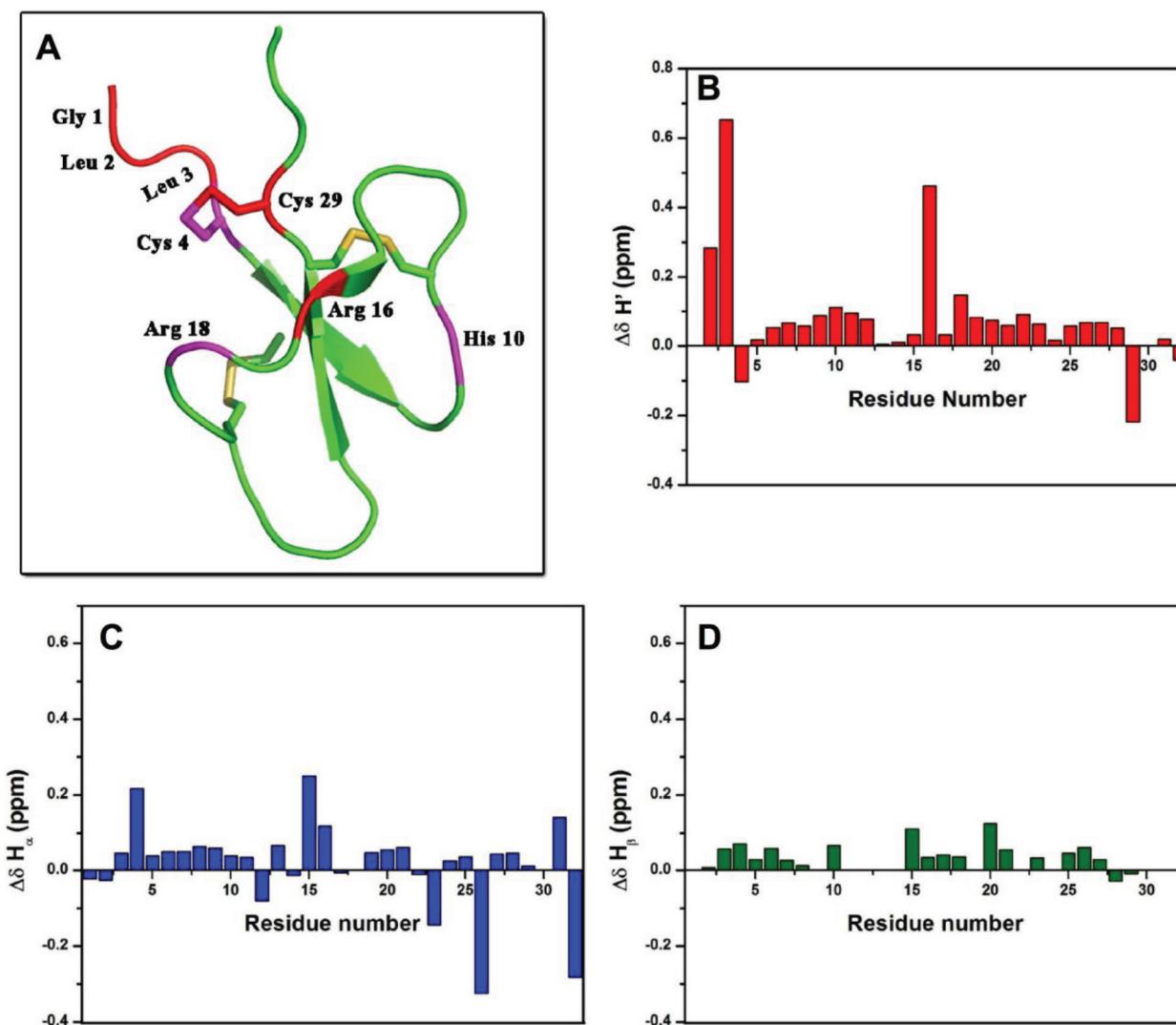


Figure 4. Chemical shifts differences of the backbone, H' and H^α , and side chain, H^β , protons between cyclic Crp4-1 and native Crp4-wt. (A) Ribbon diagram of native Crp4-wt showing the residues that exhibit a significant change in the chemical shift of the backbone amide after cyclization. Residues with a change in chemical shift larger than 0.2 ppm or between 0.1 and 0.2 ppm are colored red or magenta, respectively. Changes in the H' (B), H^α (C), and H^β (D) chemical shifts between native Crp-wt and cyclic Crp4-1 defensins reflect minimal structural perturbations due to cyclization.

of purified cyclized Crp4-1 with either free, CBD bound, or intein–CBD-bound chitin–Sepharose beads. After being extensively washed, the beads in both cases were washed with 8 M GdmCl, and the amount of eluted Crp4-1 was quantified by HPLC. The results showed that approximately $\approx 25\%$ was able to bind free or CBD-bound chitin beads, and $\approx 90\%$ of cyclized Crp4-1 was bound to intein–CBD chitin beads (data not shown). This demonstrates that although cyclized Crp4 has some weak affinity for chitin–Sepharose beads, it was bound to the beads mainly by protein–protein interactions with the gyrase intein.

Structural Characterization of Cyclized Crp4 Defensins. Because the structure of native Crp4-wt has been elucidated by 1H NMR,³⁵ we used heteronuclear NMR spectroscopy to confirm that the biosynthetic cyclized Crp4 defensins were adopting a native α -defensin fold. Given that Crp4-wt and cyclized Crp4-1 have the same sequence, we first compared the assigned backbone amide and α (H' and H^α , respectively) proton chemical shifts of cyclized Crp4-1 with those published for the native Crp4-wt peptide (Table S2 of the

Supporting Information).³⁵ As shown in Figure 4, the chemical shift differences for most of the residues were smaller than 0.1 ppm, which indicates that cyclized Crp4-1 adopts a structure that is very similar to that of native Crp4-wt. We also saw a few residues that showed chemical shift differences of >0.1 ppm. Most of these residues, however, were located in the new loop formed close to the cyclization site, including residues Leu2, Leu3, Cys4, and Cys29 (Figure 4A). Interestingly, we also noticed a few residues (His10, Arg16, and Arg18) located away from the ligation site that presented relatively large chemical shift differences (Figure 4A). The chemical shift difference associated with H' of Arg16 was ≈ 0.4 ppm; meanwhile, for the other residues, the corresponding differences were relatively smaller (≈ 0.1 ppm). On the basis of these changes, it is very likely that the hydrogen bond between the carbonyl group of Cys29 and H' of Arg16 present in the structure of native Crp4-wt is broken in cyclized Crp4-1. This change would lead to changes in the backbone dihedral angles of the neighboring amino acids. This is in agreement with the fact that the differences observed for the chemical shifts of the side chain

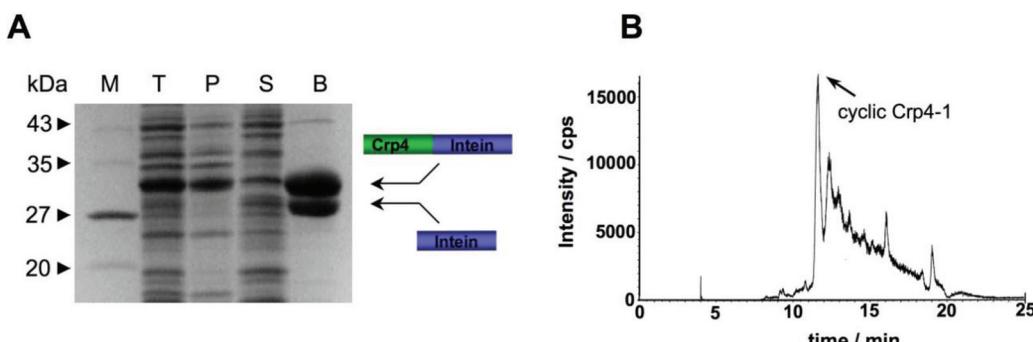


Figure 5. In vivo expression of cyclic Crp4-1 defensin in Origami2(DE3) cells. (A) SDS-PAGE analysis of the expression of Crp4–intein precursor Crp4-1. Lane M contained molecular mass protein markers, lane T total cell lysate, lane P insoluble cell lysate, lane S soluble cell lysate, and lane B affinity chromatography-purified Crp4–intein precursor. (B) Mass spectrum of the insoluble cell lysate following HPLC separation (see Experimental Procedures).

protons, H^β , are all smaller than 0.1 ppm (Figure 4D and Table S2 of the Supporting Information).

We also used heteronuclear $^1\text{H}\{^{15}\text{N}\}$ HSQC experiments to compare the structures of the different cyclized Crp4 variants. Uniformly ^{15}N -labeled cyclized Crp4 defensins were produced in vitro as described above, but *E. coli* was grown in minimal M9 medium containing $^{15}\text{NH}_4\text{Cl}$ as the only source of nitrogen. Recombinant expression of cyclized Crp4 defensins allows the introduction of NMR active isotopes (^{15}N and/or ^{13}C) in a very inexpensive fashion, thus facilitating the use of the SAR by NMR (structure–activity relationship by nuclear magnetic resonance)^{57,58} technique to study any molecular interaction between cyclized Crp4 defensins and their potential biomolecular targets. The HSQC spectra for all of the cyclized Crp4 variants were very well dispersed, indicating a well-folded structure (Figure S5 of the Supporting Information). As expected, the chemical shift differences for the backbone amide between the different cyclized Crp4 variants were also relatively small, the major differences being located in residues close to the ligation site and to Arg16 (Figure S5 of the Supporting Information). As expected, analysis of the HQSC spectrum of Crp4-2 revealed the presence of an additional Gly peak as compared to the spectrum of Crp4-1 (Figure SSB of the Supporting Information). In contrast, analysis of the HSQC spectrum of Crp4-3 revealed as many as nine additional peaks (Figure SSB of the Supporting Information). Careful integration of the intensities associated with the HSQC spectrum of cyclized Crp4-3 showed that 18 cross-peaks have intensities approximately 2 times lower than those of the rest of the peaks. These findings suggest that nine residues in cyclized Crp4-3 have two different conformations that could be exchanging at a very slow rate or not at all. We attributed the existence of two Crp4-3 conformations in solution to cis–trans isomerization of the proline in the new loop formed because most of the residues exhibiting amide peak doubling were located close to the newly formed loop.

In Vivo Biosynthesis of Cyclized Crp4 Defensins.

Encouraged by the results obtained with the in vitro GSH-induced cyclization/folding of the intein–Crp4 precursors, we decided to explore the expression of cyclized Crp4 inside *E. coli* cells. To accomplish this, we used intein precursor Crp4-1. This construct gave the best yield for the production of cyclized Crp4-1 in vitro (Table 2). The production of cyclized Crp4-1 was accomplished in Origami2(DE3) cells. These cells have mutations in the thioredoxin and glutathione reductase genes, which facilitates the formation of disulfide bonds in the

bacterial cytosol.⁵⁹ We have recently used these cells for the in vivo production of several disulfide-containing backbone-cyclized polypeptides.^{17,18,55}

The expression yield for the Crp4-1 precursor after overnight induction at room temperature with IPTG was approximately 5.5 mg/L. Under these conditions, ~35% of the Crp4–intein precursor was cleaved in vivo (Figure 5A). Using these numbers, we estimated the maximal amount of the cyclized Crp4-1 peptide that could be produced should be around 200 $\mu\text{g}/\text{L}$. Initial attempts to identify and quantify the amount of cyclized Crp4-1 in vivo by HPLC showed there to be very small amounts associated with the insoluble cell lysate and none detected in the soluble fraction. We were able to quantify the amount of cyclized Crp4-1 present in the insoluble cellular fraction using HPLC–MS/MS (Figure 5B). The yield was estimated to be $\approx 2 \mu\text{g}/\text{L}$, 100-fold lower than expected ($\approx 1\%$ yield). The cyclized Crp4-1 obtained in vivo had a mass corresponding to the folded product and coeluted with the purified product obtained in vitro, indicating that they were the same compound. The low efficiency observed for in vivo expression could be attributed to the toxicity of this defensin. α -Defensins are antimicrobial compounds that can bind and disturb the membranes of bacteria and have also been shown to inhibit the biosynthesis of peptidoglycan by binding to its precursor lipid II.^{60,61} It is likely that folded cyclized Crp4 may exert the same type of action when it is in the bacterial cytosol and could explain why Crp4-1 was found in the insoluble cell lysate. Moreover, we have shown that cyclized Crp4 defensins have affinity for the intein–CBD protein fusion, and therefore, any precursor protein present in the insoluble pellet could also bind cyclized Crp4 defensins, facilitating its immobilization to the insoluble cell lysate.

We also examined the expression of cyclized Crp4 defensins using the Crp4-4 and Crp4-5 precursors (Figure S1 of the Supporting Information), especially because the Crp4-5 precursor gave better in vivo cleavage ($\approx 70\%$). On the basis of the expression levels, the predicted yield for Crp4-4 is approximately 20 $\mu\text{g}/\text{L}$ and for Crp4-5 almost 200 $\mu\text{g}/\text{L}$. In both cases, the corresponding folded cyclized Crp4 variant was found in the insoluble cell lysate at yields similar to that found for Crp4-1 (1–2 $\mu\text{g}/\text{L}$). These three precursors gave very different yields for cyclization/folding in vitro. The fact that they provide a similar yield when expressed in vivo may suggest that the production in vivo could be limited by the cellular activity of the defensin.

Antimicrobial and Hemolytic Activities of Cyclized Crp4 Defensins. Native Crp4-wt has potent antimicrobial activities against a broad spectrum of microorganisms, including Gram-positive and Gram-negative bacteria.^{42,46,62} To explore the effect of cyclization on its biological activity, we tested the antimicrobial activities of the cyclized Crp4 defensins against several Gram-positive and Gram-negative bacteria. As shown in Figure 6, all cyclized defensins exhibited a dose-dependent

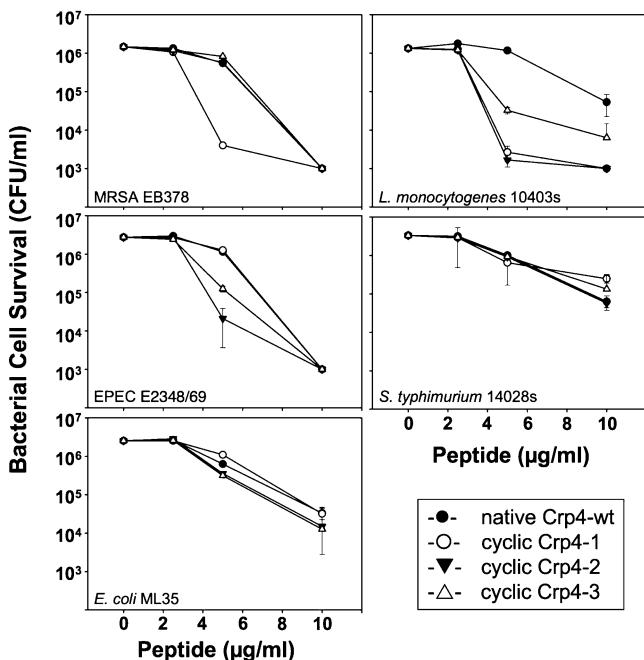


Figure 6. Bactericidal activities of Crp4-wt and cyclic Crp4 variants. Exponentially growing MRSA, *L. monocytogenes* 10403s, EPEC E2348/69, *S. typhimurium* 14028s, and *E. coli* ML35 were exposed to peptides at 37 °C in 50 μL of PIPES-TSB buffer for 1 h (see Experimental Procedures). Following peptide exposure, the bacteria were plated on TSB-agar plates and incubated overnight at 37 °C. Surviving bacteria were counted as CFU at each peptide concentration, and count values below 1 × 10³ CFU/mL signify that no colonies were detected.

killing of the bacteria tested, including methicillin resistant *Staphylococcus aureus* (MRSA) strain EB378 and enteropathogenic *E. coli* (EPEC) strain E2348/69. Overall, the cyclic peptides were most active against *Listeria monocytogenes* 10403s, killing 90–100% at 10 μg/mL, and were least active against *Salmonella typhimurium* 14028s but still killed approximately 90% of the bacteria at 10 μg/mL. The cyclic Crp4 peptides exhibited bactericidal activity equivalent to or in some cases greater than that of native Crp4-wt (Figure 6). For example, cyclized Crp4-1 and Crp4-2 were more potent than native Crp4-wt against *L. monocytogenes* 10403s at 5 and 10 μg/mL. Similar results were obtained for cyclized Crp4-1 and Crp4-2 against MRSA and EPEC, respectively, at 5 mg/mL. These data demonstrate that cyclization preserves the biological activity of Crp4 and, for some bacterial cell targets, improves bactericidal activity.

Analysis of the antimicrobial activities in the presence of increasing NaCl concentrations showed minor differences at 160 mM NaCl, in which 10 μg/mL native Crp4-wt killed 70% and cyclized Crp4-1 killed 80% of *L. monocytogenes* 10403s and *E. coli* ML35 (Figure S6a,b of the Supporting Information). These data contrast with previous studies of cyclic rabbit NP-1,

which was more active than natural NP-1 against *E. coli* in the presence of 100 mM NaCl.³⁷ This may be due to differences in the primary structures of Crp4 versus NP-1 (Figure 1), despite their highly similar three-dimensional folds. The antimicrobial activities of Crp4-wt and cyclic variants in the presence of 5% heat-inactivated human serum were attenuated, but 75–95% of EPEC exposed to 10 μg/mL peptide were still killed (Figure S6C of the Supporting Information). These data suggest the cyclic Crp4 variants retain antimicrobial activity under conditions that mimic an in vivo environment and, therefore, have the potential to be developed as a drug lead, particularly against antibiotic resistant bacteria.

To assess the selectivity of Crp4 and cyclized Crp4 defensins, we tested the cytotoxicity of native Crp4-wt and cyclized Crp4-1 in a hemolysis assay against human RBCs. Both peptides lacked hemolytic activity at concentrations up to 100 μg/mL, while the positive control peptide, melittin, gave approximately 75% hemolysis at the same concentration (Table 3). Additionally,

Table 3. Hemolytic Activities of Native and Cyclized Crp4 Defensins^a

peptide	0 μg/ mL	1 μg/ mL	12.5 μg/ mL	25 μg/ mL	50 μg/ mL	100 μg/ mL
Crp4-wt	4.2%	n/d ^b	4.1%	3.4%	3.0%	3.1%
Crp4-1	4.2%	n/d ^b	5.4%	5.8%	5.9%	4.9%
melittin	4.2%	3.6%	n/d ^b	n/d ^b	n/d ^b	74.6%

^aThe peptide melittin was used as a control. ^bNot determined.

native Crp4 has been found to be noncytotoxic to the mouse macrophage cell line, RAW 264.7 (unpublished work from the lab of A. J. Ouellette). These data demonstrate native and cyclic Crp4 are selective against bacteria.

Stability of Crp4 in Human Serum. Proteomics analyses of mouse colonic luminal contents have shown that intact, active α -defensins persist after secretion by Paneth cells of the small intestine.⁶³ Cryptdin peptides have been recovered from washing of mouse jejunum and ileum^{42,47} and the distal colonic lumen,⁶³ demonstrating their inherent resistance to proteolysis in the gastrointestinal environment conferred by the disulfide array. The stability of Crp4 was further assessed by incubation of native and cyclic peptides in 100% human serum, a location where Crp4 does not naturally occur (Figure 7). Both the native and cyclic peptides remained intact, with no hydrolysis of the N- and C-termini of native Crp4 (data not shown). In contrast, degradation of a disulfide-null mutant of native Crp4, in which the six Cys residues are mutated to Ala (Crp4-6C/A), and an S-carboxamidomethylated linear Crp4 (Crp4-R/A) (Table 1) began after just 2 min in human serum, and they were completely proteolyzed within 10–30 min (Figure 7, inset). Although Crp4-6C/A maintains potent antimicrobial activity in vitro, it is susceptible to degradation by matrix metalloproteinase (MMP) 7, the mouse pro- α -defensin convertase.^{50,64} These stability results open the possibility of developing Crp4 as a therapeutic agent. Although native Crp4 was stable in serum for 48 h, it is possible cyclized Crp4 variants may be more stable at longer time points or in vivo, when injected into or fed to animals.

DISCUSSION

Defensins are innate immune peptides that play an important role in the host defense of mammals. In this study, we have produced several backbone-cyclized variants of a mouse

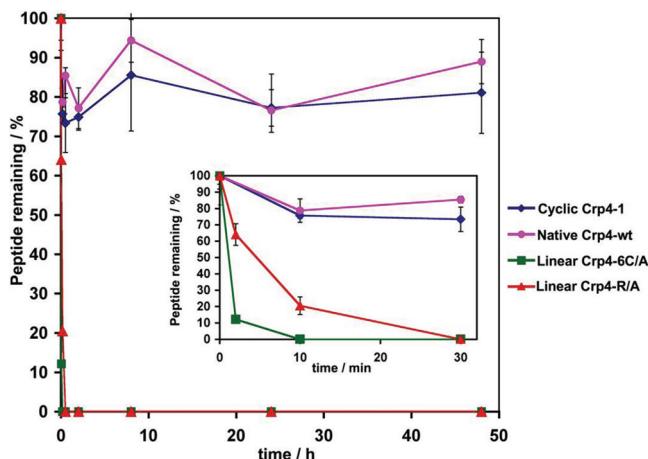


Figure 7. Stability of Crp4 and variants in human serum. Peptides at 100 μ g/mL were incubated with heat-inactivated serum at 37 °C and analyzed by C4 RP-HPLC. Linear Crp4-6C/A is a native Crp4-wt variant in which the six Cys are mutated to Ala. Linear Crp4-R/A corresponds to a linear Crp4-1 construct in which the six Cys residues are alkylated with iodoacetamide.

α -defensin in *E. coli* cells using modified protein splicing units. The cyclized peptides were characterized by NMR, antimicrobial activity, and serum stability.

The different cyclized versions of Crp4 were produced recombinantly either in vitro or in vivo. The best expression yield was obtained for Crp4-1 when it was cyclized in vitro (\approx 200 μ g/L). The yields for the other cyclized Crp4 variants ranged from 170 to 35 μ g/L (Table 2). In vivo expression was by far less efficient with yields of \approx 2 μ g/L. All of the examined cyclized Crp4 variants gave a similar yield when expressed in vivo that was independent of the precursor used. This strongly suggests that the amount of cyclized defensins that could be produced in vivo may be limited by the intracellular toxicity of the defensin. α -Defensins are antimicrobial compounds that can bind and disturb the membranes of bacteria and have also been shown to inhibit the biosynthesis of peptidoglycan by binding to its precursor lipid II.^{60,61} It is possible that folded cyclized Crp4 may exert the same type of action when it is in the bacterial cytosol, explaining the intracellular toxicity. Our data also indicate that cyclized Crp4 variants bind strongly to the intein–CBD fusion protein, specifically to the intein region. At this time, it is unknown how the peptide is binding to the protein. It may involve electrostatic interactions with the intein because Crp4 is highly cationic and the gyrase intein has an overall anionic charge. Interestingly, in vivo production of cyclized Crp4 defensins always provided the peptide on the insoluble fraction; therefore, it is possible that this occurs through binding to the insoluble intein–CBD protein. It may also be bound to insoluble bacterial membranes or cell wall precursors such as lipid II. Additional experiments to fractionate the pellet are necessary to delineate these results.

Structural characterization of the cyclized Crp4 variants by NMR revealed that they adopt structures that are very similar to that of native Crp4. This indicates that backbone cyclization did not significantly change the native fold of Crp4. The antimicrobial assays also demonstrated that backbone cyclization did not affect the biological activities of the peptide. In fact, some of the cyclized defensins exhibited better bactericidal activity (Figure 6). Interestingly, native and cyclic Crp4 showed no hemolytic activity against human red blood

cells, demonstrating their selectivity against bacteria. Additionally, Crp4-wt and cyclic Crp4-1 defensins both demonstrated high stability to human serum, with half-lives of >48 h. In contrast, mutation or alkylation of the Cys residues resulted in complete degradation of the corresponding linear peptides by human serum in <30 min. On the basis of these data, it is likely the three disulfide bonds are essential components in the stabilization of the peptide structures. The presence of the disulfide array has previously been shown to be a necessary component in resistance to proteolytic degradation but does not affect antimicrobial activity.⁵⁰ Altogether, these properties make cyclized Crp4 defensins promising scaffolds for drug development of novel antibiotics, although further studies may be required to evaluate their metabolic stability and bioavailability.

Novel antimicrobial agents are necessary to overcome the threat of prevalent antibiotic resistant pathogens. On the basis of the data reported here, this peptide can potentially be used as a stable scaffold to generate more enhanced antimicrobial drugs. Although a library of Crp4 sequences is not useful in bacterial expression systems, they have the potential to be expressed in yeast or mammalian cells providing there is no intracellular cytotoxicity. This may be useful for in vivo screening against intracellular pathogens, including viruses and parasites.⁶⁵ Studies have shown that antimicrobial peptides can decrease the viability of intracellular *Mycobacterium tuberculosis*⁶⁶ and inhibit the proliferation of intracellular *L. monocytogenes* in macrophages;⁶⁷ therefore, the expression of defensin libraries in mammalian cells is a promising method for developing and screening for more effective antimicrobials against intracellular pathogens.

ASSOCIATED CONTENT

Supporting Information

Supplementary results mentioned in the text, including two tables and seven figures. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

*Department of Pharmacology and Pharmaceutical Sciences, School of Pharmacy, University of Southern California, 1985 Zonal Ave., PSC 616, Los Angeles, CA 90033. Phone: (323) 442-1417. E-mail: camarej@usc.edu.

Funding

This work was supported by National Institutes of Health (NIH) Research Grants R01-GM090323 (J.A.C.) and DK044632 and AI059346 (A.J.O.), and NIH Grant SR01GM085006-02 (A.S.) and by Department of Defense Congressionally Directed Medical Research Program Grant PC09305 (J.A.C.).

ACKNOWLEDGMENTS

We thank Caroline Martel and Sabine Chauveau for technical assistance. We thank Dr. Annie Wong-Beringer (University of Southern California) for generously providing MRSA clinical isolate EB378 and Dr. Gail Hecht (University of Illinois at Chicago, Chicago, IL) for EPEC strain E2348/69.

ABBREVIATIONS

CBD, chitin binding domain; CFU, colony-forming unit; Crp4, cryptdin 4; DPBS, Dulbecco's phosphate-buffered saline;

EDTA, ethylenediaminetetraacetic acid; EPEC, enteropathogenic *E. coli*; GdmCl, guanidinium chloride; GSH, reduced glutathione; HPLC, high-performance liquid chromatography; HSQC, heteronuclear single-quantum coherence; IPTG, isopropyl β -D-1-thiogalactopyranoside; NCL, native chemical ligation; NMR, nuclear magnetic resonance; MRSA, methicillin resistant *St. aureus*; NOESY, nuclear Overhauser effect spectroscopy; PIPES, 1,4-piperazinediethanesulfonic acid; PMSF, phenylmethanesulfonyl fluoride; RBCs, red blood cells; SFTI-1, sunflower trypsin inhibitor 1; TFA, trifluoroacetic acid; TOCSY, total correlation spectroscopy; TSP, trypticase soy broth; UV, ultraviolet.

■ REFERENCES

- Unger, T., Oren, Z., and Shai, Y. (2001) The effect of cyclization of magainin 2 and melittin analogues on structure, function, and model membrane interactions: Implication to their mode of action. *Biochemistry* 40, 6388–6397.
- Dathe, M., Nikolenko, H., Klose, J., and Bienert, M. (2004) Cyclization increases the antimicrobial activity and selectivity of arginine- and tryptophan-containing hexapeptides. *Biochemistry* 43, 9140–9150.
- Clark, R. J., Fischer, H., Dempster, L., Daly, N. L., Rosengren, K. J., Nevin, S. T., Meunier, F. A., Adams, D. J., and Craik, D. J. (2005) Engineering stable peptide toxins by means of backbone cyclization: Stabilization of the α -conotoxin MII. *Proc. Natl. Acad. Sci. U.S.A.* 102, 13767–13772.
- Adessi, C., and Soto, C. (2002) Converting a peptide into a drug: Strategies to improve stability and bioavailability. *Curr. Med. Chem.* 9, 963–978.
- Camarero, J. A., Fushman, D., Sato, S., Giri, I., Cowburn, D., Raleigh, D. P., and Muir, T. W. (2001) Rescuing a destabilized protein fold through backbone cyclization. *J. Mol. Biol.* 308, 1045–1062.
- Clark, R. J., Jensen, J., Nevin, S. T., Callaghan, B. P., Adams, D. J., and Craik, D. J. (2010) The engineering of an orally active conotoxin for the treatment of neuropathic pain. *Angew. Chem., Int. Ed.* 49, 6545–6548.
- Trabi, M., and Craik, D. J. (2002) Circular proteins: No end in sight. *Trends Biochem. Sci.* 27, 132–138.
- Camarero, J. A., and Muir, T. W. (1997) Chemoselective backbone cyclization of unprotected peptides. *Chem. Commun.*, 1369–1370.
- Camarero, J. A., Cotton, G. J., Adeva, A., and Muir, T. W. (1998) Chemical ligation of unprotected peptides directly from a solid support. *J. Pept. Res.* 51, 303–316.
- Camarero, J. A., Pavel, J., and Muir, T. W. (1998) Chemical Synthesis of a Circular Protein Domain: Evidence for Folding-Assisted Cyclization. *Angew. Chem., Int. Ed.* 37, 347–349.
- Shao, Y., Lu, W. Y., and Kent, S. B. H. (1998) A novel method to synthesize cyclic peptides. *Tetrahedron Lett.* 39, 3911–3914.
- Tam, J. P., and Lu, Y. A. (1998) A biomimetic strategy in the synthesis and fragmentation of cyclic protein. *Protein Sci.* 7, 1583–1592.
- Camarero, J. A., and Mitchell, A. R. (2005) Synthesis of proteins by native chemical ligation using Fmoc-based chemistry. *Protein Pept. Lett.* 12, 723–728.
- Camarero, J. A., and Muir, T. W. (1999) Biosynthesis of a Head-to-Tail Cyclized Protein with Improved Biological Activity. *J. Am. Chem. Soc.* 121, 5597–5598.
- Evans, T. C., Benner, J., and Xu, M.-Q. (1999) The cyclization and polymerization of bacterially expressed proteins using modified sef-splicing inteins. *J. Biol. Chem.* 274, 18359–18381.
- Scott, C. P., Abel-Santos, E., Wall, M., Wahnon, D., and Benkovic, S. J. (1999) Production of cyclic peptides and proteins in vivo. *Proc. Natl. Acad. Sci. U.S.A.* 96, 13638–13643.
- Camarero, J. A., Kimura, R. H., Woo, Y. H., Shekhtman, A., and Cantor, J. (2007) Biosynthesis of a fully functional cyclotide inside living bacterial cells. *ChemBioChem* 8, 1363–1366.
- Austin, J., Wang, W., Puttamatappa, S., Shekhtman, A., and Camarero, J. A. (2009) Biosynthesis and biological screening of a genetically encoded library based on the cyclotide MCoTI-I. *ChemBioChem* 10, 2663–2670.
- Austin, J., Kimura, R. H., Woo, Y. H., and Camarero, J. A. (2009) In vivo biosynthesis of an Ala-scan library based on the cyclic peptide SFTI-1. *Amino Acids* 38, 1313–1322.
- Selsted, M. E., Harwig, S. S., Ganz, T., Schilling, J. W., and Lehrer, R. I. (1985) Primary structures of three human neutrophil defensins. *J. Clin. Invest.* 76, 1436–1439.
- Ouellette, A. J., Miller, S. I., Henschchen, A. H., and Selsted, M. E. (1992) Purification and primary structure of murine cryptdin-1, a Paneth cell defensin. *FEBS Lett.* 304, 146–148.
- Tang, Y. Q., Yuan, J., Osapay, G., Osapay, K., Tran, D., Miller, C. J., Ouellette, A. J., and Selsted, M. E. (1999) A cyclic antimicrobial peptide produced in primate leukocytes by the ligation of two truncated α -defensins. *Science* 286, 498–502.
- Tanabe, H., Yuan, J., Zaragoza, M. M., Dandekar, S., Henschchen-Edman, A., Selsted, M. E., and Ouellette, A. J. (2004) Paneth cell α -defensins from rhesus macaque small intestine. *Infect. Immun.* 72, 1470–1478.
- Selsted, M. E., and Ouellette, A. J. (2005) Mammalian defensins in the antimicrobial immune response. *Nat. Immunol.* 6, 551–557.
- Niyonsaba, F., Ushio, H., Nakano, N., Ng, W., Sayama, K., Hashimoto, K., Nagaoka, I., Okumura, K., and Ogawa, H. (2007) Antimicrobial peptides human β -defensins stimulate epidermal keratinocyte migration, proliferation and production of proinflammatory cytokines and chemokines. *J. Invest. Dermatol.* 127, 594–604.
- Steinstraesser, L., Koehler, T., Jacobsen, F., Daigeler, A., Goertz, O., Langer, S., Kesting, M., Steinau, H., Eriksson, E., and Hirsch, T. (2008) Host defense peptides in wound healing. *Mol. Med.* 14, 528–537.
- Yang, D., Chen, Q., Chertov, O., and Oppenheim, J. J. (2000) Human neutrophil defensins selectively chemoattract naive T and immature dendritic cells. *J. Leukocyte Biol.* 68, 9–14.
- Territo, M. C., Ganz, T., Selsted, M. E., and Lehrer, R. (1989) Monocyte-chemotactic activity of defensins from human neutrophils. *J. Clin. Invest.* 84, 2017–2020.
- Yang, D., Liu, Z. H., Tewary, P., Chen, Q., de la Rosa, G., and Oppenheim, J. J. (2007) Defensin participation in innate and adaptive immunity. *Curr. Pharm. Des.* 13, 3131–3139.
- Rehaume, L. M., and Hancock, R. E. (2008) Neutrophil-derived defensins as modulators of innate immune function. *Crit. Rev. Immunol.* 28, 185–200.
- Bhattacharjya, S. (2010) De novo designed lipopolysaccharide binding peptides: Structure based development of antiendotoxic and antimicrobial drugs. *Curr. Med. Chem.* 17, 3080–3093.
- Scott, M. G., Vreugdenhil, A. C., Buurman, W. A., Hancock, R. E., and Gold, M. R. (2000) Cutting edge: Cationic antimicrobial peptides block the binding of lipopolysaccharide (LPS) to LPS binding protein. *J. Immunol.* 164, 549–553.
- Motzkus, D., Schulz-Maronde, S., Heitland, A., Schulz, A., Forssmann, W. G., Jubner, M., and Maronde, E. (2006) The novel β -defensin DEFB123 prevents lipopolysaccharide-mediated effects in vitro and in vivo. *FASEB J.* 20, 1701–1702.
- Droin, N., Hendra, J. B., Ducoroy, P., and Solaray, E. (2009) Human defensins as cancer biomarkers and antitumour molecules. *J. Proteomics* 72, 918–927.
- Rosengren, K. J., Daly, N. L., Fornander, L. M., Jonsson, L. M., Shirafuji, Y., Qu, X., Vogel, H. J., Ouellette, A. J., and Craik, D. J. (2006) Structural and functional characterization of the conserved salt bridge in mammalian paneth cell α -defensins: Solution structures of mouse CRYPTDIN-4 and (E15D)-CRYPTDIN-4. *J. Biol. Chem.* 281, 28068–28078.
- Bals, R., Goldman, M. J., and Wilson, J. M. (1998) Mouse β -defensin 1 is a salt-sensitive antimicrobial peptide present in epithelia of the lung and urogenital tract. *Infect. Immun.* 66, 1225–1232.

(37) Yu, Q., Lehrer, R. I., and Tam, J. P. (2000) Engineered salt-insensitive α -defensins with end-to-end circularized structures. *J. Biol. Chem.* 275, 3943–3949.

(38) Ayabe, T., Satchell, D. P., Wilson, C. L., Parks, W. C., Selsted, M. E., and Ouellette, A. J. (2000) Secretion of microbicidal α -defensins by intestinal Paneth cells in response to bacteria. *Nat. Immunol.* 1, 113–118.

(39) Ouellette, A. J., Greco, R. M., James, M., Frederick, D., Naftilan, J., and Fallon, J. T. (1989) Developmental regulation of cryptdin, a corticostatin/defensin precursor mRNA in mouse small intestinal crypt epithelium. *J. Cell Biol.* 108, 1687–1695.

(40) Eisenhauer, P. B., Harwig, S. S., and Lehrer, R. I. (1992) Cryptdins: Antimicrobial defensins of the murine small intestine. *Infect. Immun.* 60, 3556–3565.

(41) Ouellette, A. J., Darmoul, D., Tran, D., Huttner, K. M., Yuan, J., and Selsted, M. E. (1999) Peptide localization and gene structure of cryptdin 4, a differentially expressed mouse Paneth cell α -defensin. *Infect. Immun.* 67, 6643–6651.

(42) Ouellette, A. J., Satchell, D. P., Hsieh, M. M., Hagen, S. J., and Selsted, M. E. (2000) Characterization of luminal Paneth cell α -defensins in mouse small intestine. Attenuated antimicrobial activities of peptides with truncated amino termini. *J. Biol. Chem.* 275, 33969–33973.

(43) Salzman, N. H., Hung, K., Haribhai, D., Chu, H., Karlsson-Sjoberg, J., Amir, E., Teggatz, P., Barman, M., Hayward, M., Eastwood, D., Stoel, M., Zhou, Y., Sodergren, E., Weinstock, G. M., Bevins, C. L., Williams, C. B., and Bos, N. A. (2010) Enteric defensins are essential regulators of intestinal microbial ecology. *Nat. Immunol.* 11, 76–83.

(44) Satoh, Y. (1988) Effect of live and heat-killed bacteria on the secretory activity of Paneth cells in germ-free mice. *Cell Tissue Res.* 251, 87–93.

(45) Satoh, Y., Ishikawa, K., Oomori, Y., Yamano, M., and Ono, K. (1989) Effects of cholecystokinin and carbamylcholine on Paneth cell secretion in mice: A comparison with pancreatic acinar cells. *Anat. Rec.* 225, 124–132.

(46) Ouellette, A. J., Hsieh, M. M., Nosek, M. T., Cano-Gauci, D. F., Huttner, K. M., Buick, R. N., and Selsted, M. E. (1994) Mouse Paneth cell defensins: Primary structures and antibacterial activities of numerous cryptdin isoforms. *Infect. Immun.* 62, 5040–5047.

(47) Selsted, M. E., Miller, S. I., Henschen, A. H., and Ouellette, A. J. (1992) Enteric defensins: Antibiotic peptide components of intestinal host defense. *J. Cell Biol.* 118, 929–936.

(48) Sancheti, H., and Camarero, J. A. (2009) “Splicing up” drug discovery. Cell-based expression and screening of genetically-encoded libraries of backbone-cyclized polypeptides. *Adv. Drug Delivery Rev.* 61, 908–917.

(49) Kimura, R. H., Tran, A. T., and Camarero, J. A. (2006) Biosynthesis of the cyclotide kalata B1 by using protein splicing. *Angew. Chem., Int. Ed.* 45, 973–976.

(50) Maemoto, A., Qu, X., Rosengren, K. J., Tanabe, H., Henschen-Edman, A., Craik, D. J., and Ouellette, A. J. (2004) Functional analysis of the α -defensin disulfide array in mouse cryptdin-4. *J. Biol. Chem.* 279, 44188–44196.

(51) Figueiredo, S., Mastroianni, J. R., Tai, K. P., and Ouellette, A. J. (2010) Expression and purification of recombinant α -defensins and α -defensin precursors in *Escherichia coli*. *Methods Mol. Biol.* 618, 47–60.

(52) Puttamadappa, S. S., Jagadish, K., Shekhtman, A., and Camarero, J. A. (2010) Backbone Dynamics of Cyclotide MCoTI-I Free and Complexed with Trypsin. *Angew. Chem., Int. Ed.* 49, 7030–7034.

(53) Cavanagh, J., and Rance, M. (1992) Suppression of cross relaxation effects in TOCSY spectra via a modified DISI-2 mixing sequence. *J. Magn. Reson.* 96, 670–678.

(54) Wuthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.

(55) Austin, J., Kimura, R. H., Woo, Y. H., and Camarero, J. A. (2010) In vivo biosynthesis of an Ala-scan library based on the cyclic peptide SFTI-1. *Amino Acids* 38, 1313–1322.

(56) Chou, P. Y., and Fasman, G. D. (1977) β -Turns in proteins. *J. Mol. Biol.* 115, 135–175.

(57) Shuker, S. B., Hajduk, P. J., Meadows, R. P., and Fesik, S. W. (1996) Discovering high-affinity ligands for proteins: SAR by NMR. *Science* 274, 1531–1534.

(58) Hajduk, P. J., Meadows, R. P., and Fesik, S. W. (1997) Discovering high-affinity ligands for proteins. *Science* 278, 497–499.

(59) Bessette, P. H., Aslund, F., Beckwith, J., and Georgiou, G. (1999) Efficient folding of proteins with multiple disulfide bonds in the *Escherichia coli* cytoplasm. *Proc. Natl. Acad. Sci. U.S.A.* 96, 13703–13708.

(60) de Leeuw, E., Li, C., Zeng, P., Diepeveen-de Buin, M., Lu, W. Y., Breukink, E., and Lu, W. (2010) Functional interaction of human neutrophil peptide-1 with the cell wall precursor lipid II. *FEBS Lett.* 584, 1543–1548.

(61) Wilmes, M., Cammue, B. P., Sahl, H. G., and Thevissen, K. (2011) Antibiotic activities of host defense peptides: More to it than lipid bilayer perturbation. *Nat. Prod. Rep.* 28, 1350–1358.

(62) Harwig, S. S., Eisenhauer, P. B., Chen, N. P., and Lehrer, R. I. (1995) Cryptdins: Endogenous antibiotic peptides of small intestinal Paneth cells. *Adv. Exp. Med. Biol.* 371A, 251–255.

(63) Mastroianni, J. R., and Ouellette, A. J. (2009) α -Defensins in enteric innate immunity: Functional Paneth cell α -defensins in mouse colonic lumen. *J. Biol. Chem.* 284, 27848–27856.

(64) Ouellette, A. J. Paneth cell α -defensins in enteric innate immunity. *Cell. Mol. Life Sci.* 68, 2215–2229.

(65) Foureau, D. M., Mielcarz, D. W., Menard, L. C., Schulthess, J., Werts, C., Vasseur, V., Ryffel, B., Kasper, L. H., and Buzoni-Gatel, D. (2010) TLR9-dependent induction of intestinal α -defensins by *Toxoplasma gondii*. *J. Immunol.* 184, 7022–7029.

(66) Tan, B. H., Meinken, C., Bastian, M., Bruns, H., Legaspi, A., Ochoa, M. T., Krutzik, S. R., Bloom, B. R., Ganz, T., Modlin, R. L., and Stenger, S. (2006) Macrophages acquire neutrophil granules for antimicrobial activity against intracellular pathogens. *J. Immunol.* 177, 1864–1871.

(67) Arnett, E., Lehrer, R. I., Pratikhya, P., Lu, W., and Seveau, S. (2011) Defensins enable macrophages to inhibit the intracellular proliferation of *Listeria monocytogenes*. *Cell. Microbiol.* 13, 635–651.

Recombinant production of rhesus θ -defensin-1 (RTD-1) using a bacterial expression system[†]

Andrew Gould,^{‡,a} Yilong Li,^{‡,a} Subhabrata Majumder,^b Angie E. Garcia,^a Patrick Carlsson,^a Alexander Shekhtman^b and Julio A. Camarero^{*ac}

Received 1st November 2011, Accepted 19th January 2012

DOI: 10.1039/c2mb05451e

Defensins are antimicrobial peptides that are important in the innate immune defense of mammals. In contrast to mammalian α - and β -defensins, rhesus θ -defensin-1 (RTD-1) comprises only 18 amino acids stabilized by three disulfide bonds and an unusual backbone cyclic topology. In this work we report for the first time the recombinant expression of the fully folded θ -defensin RTD-1 using a bacterial expression system. This was accomplished using an intramolecular native chemical ligation in combination with a modified protein-splicing unit. RTD-1 was produced either *in vitro* or *in vivo*. In-cell production of RTD-1 was estimated to reach an intracellular concentration of $\sim 4 \mu\text{M}$. Recombinant RTD-1 was shown to be correctly folded as characterized by heteronuclear-NMR and by its ability to specifically inhibit lethal factor protease. The recombinant production of folded θ -defensins opens the possibility to produce peptide libraries based on this peptide scaffold that could be used to develop in-cell screening and directed evolution technologies.

Introduction

Defensins are cysteine-rich antimicrobial peptides that are important in the innate immune defense of mammals.^{1–3} They are classically known for their antimicrobial activities, but they are also involved in other defense mechanisms including wound healing, immune modulation, neutralization of endotoxin, and anti-cancer activities.^{3,4} Mammalian defensins are cationic peptides with largely β -sheet structures and six conserved cysteines. They can be classified into three structurally distinct groups, α -, β - and θ -defensins. The overall fold of α - and β -defensins is quite similar despite differences in disulfide connectivities, and the presence of an N-terminal α -helix segment in β -defensins that is missing in α -defensins.⁵ θ -Defensins, on the other hand, are backbone cyclized peptides formed by the head-to-tail covalent assembly of two nonapeptides derived from α -defensin related precursors,¹ and to date, are the only known cyclic polypeptides expressed in animals.¹

Rhesus θ -defensin-1 (RTD-1) was the first θ -defensin to be discovered from an extract of leukocytes from Rhesus macaques,¹

and in contrast with α - and β -defensins was shown to have a β -hairpin-like structure with two anti-parallel β -strands stabilized by three disulfides in a ladder configuration (Fig. 1A).⁶ Since then, other less abundant RTD variants, named RTD-2 to RTD-6, were also found in Rhesus macaques.^{7–9} Circular θ -defensins have also been isolated from other primate species.^{10–12} Interestingly, humans possess genes encoding θ -defensins, but they have lost the ability to produce the peptides due to a stop codon mutation within the signal sequence that prevents subsequent translation.¹³

θ -Defensins have both Gram-positive and Gram-negative antibacterial activity,¹ although this activity strongly depends on the buffer conditions used in the assays.¹⁴ For example, the antimicrobial activity of RTD-1 is negatively affected by the presence of 10% human serum.¹⁴ θ -Defensins also have anti-fungal¹ and anti-HIV^{13,15} activities. Chemically-synthesized θ -defensins (called retrocyclins), which are derived from the human pseudogene sequences, have been shown to protect human cells from infection by HIV-1¹³ and have been evaluated as a topical anti-HIV agent for the prevention of HIV transmission,^{16–18} showing promise when compared to other topical anti-HIV drugs in pre-clinical development.¹⁵ It is likely that the ability of θ -defensins to bind gp120 and CD4 glycoproteins is integrally related to their ability to protect cells from HIV-1 infection.¹⁹ θ -Defensins have also been shown to inactivate germinating anthrax spores and act as a competitive inhibitor of anthrax lethal factor (LF) protease.²⁰

Although the precursor genes associated with θ -defensins have been identified,¹ the biochemical mechanism responsible for their post-translational biosynthesis has not been elucidated yet. Our group has recently developed a method for the biosynthesis of

^a Department of Pharmacology and Pharmaceutical Sciences, University of Southern California, 1985 Zonal Avenue, PSC 616, Los Angeles, CA 90033, USA. E-mail: camarej@usc.edu; Tel: +1 323-442-1417

^b Department of Chemistry, State University of New York, Albany, NY 12222, USA

^c Department of Chemistry, University of Southern California, Los Angeles, CA 90033, USA

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c2mb05451e

‡ These authors contributed equally to this work.

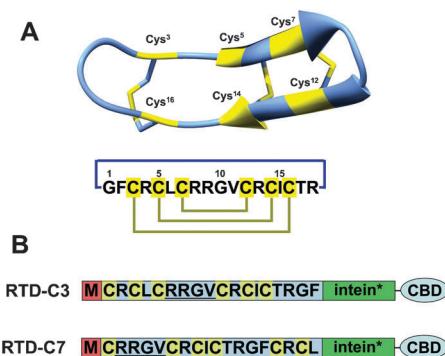


Fig. 1 (A) Primary and tertiary structure of rhesus θ -defensin 1 (RTD-1) (PDB ID code: 1HVZ).⁶ The backbone cyclized peptide (connecting bond shown in blue) is stabilized by the three disulfide bonds in a ladder formation (disulfide bonds shown in yellow). (B) Design of the two RTD-intein precursors used in this work, RTD-C3 and RTD-C7. The precursors consist of an RTD-1 based linear peptide attached to an engineered *Mxe* GyrA intein (represented by an asterisk) and the chitin binding domain (CBD). The sequence of this intein is available at <http://tools.neb.com/inbase/intein.php?name=Mxe+GyrA>. An N-terminal Met was added right in front of the required N-terminal Cys residue. The N-terminal Met residue is removed *in vivo* by endogenous methionine aminopeptidases (AMPs) producing an N-terminal Cys available for cyclization through intramolecular native chemical ligation. The sequence RRGV is underlined for reference.

backbone cyclized peptides using intramolecular Native Chemical Ligation (NCL) in combination with a modified protein-splicing unit or intein.^{21,22} This process requires the presence, within the same polypeptide sequence, of an N-terminal Cys residue and a C-terminal α -thioester function.^{23,24} We have successfully used this approach for the recombinant production of several naturally occurring backbone cyclized polypeptides using standard bacterial expression systems.^{25,26} Encouraged by these results we decided to explore the potential of this approach for the biological production of folded θ -defensins using bacterial expression systems.

Using RTD-1 as a model system, we show here that folded backbone cyclized θ -defensins can be produced recombinantly using bacterial expression systems. Folded RTD-1 can be either produced *in vitro* or *in vivo* with similar yields. In-cell production of RTD-1 can reach intracellular concentrations of $\sim 4 \mu\text{M}$. Recombinant RTD-1 was shown to adopt a native folded structure as determined by heteronuclear NMR and was fully active as shown by inhibition of anthrax LF protease. Recombinant expression of θ -defensins makes it possible to introduce NMR active isotopes (^{15}N and/or ^{13}C) in a very inexpensive fashion, thus facilitating the use of NMR to study any molecular interaction between θ -defensins and their potential biomolecular targets. These results also open the intriguing possibility for in-cell production of genetically-encoded peptide libraries based on this peptide scaffold that could be used to develop in-cell screening and directed evolution technologies.

Results and discussion

The recombinant expression of RTD-1 was carried out by using a modified protein splicing unit to assist the intramolecular native chemical ligation (NCL) required for backbone cyclization.^{23,27}

RTD-1 has six Cys residues that may be used for cyclization. In order to facilitate the cyclization reaction, we decided to use the Cys residues located in positions 3 and 7 (Fig. 1). These Cys residues are the only ones in the sequence of RTD-1 that do not have a charged or β -branched residue N-terminally adjacent, which should facilitate the kinetics of the cyclization reaction without affecting the splicing activity of the intein. Accordingly, two different RTD-1 linear precursors (RTD-C3 and RTD-C7) were cloned in frame with a modified *Mxe* GyrA intein (Fig. 1B). The N-terminal Met residue in both constructs is efficiently removed by endogenous methionine aminopeptidases when expressed in *Escherichia coli*^{21,22} therefore yielding the required N-terminal Cys for intramolecular NCL. At the same time, the modified intein allows the generation of the required α -thioester at the junction between the C-terminal end of RTD-1 and the intein (Fig. 1).

Both RTD-1 intein fusion protein precursors (RTD-C3 and RTD-C7) were expressed in *E. coli* BL21(DE3) cells at 30 °C for 3 h, and purified by affinity chromatography using chitin-sepharose beads. The RTD-intein precursors have a chitin binding domain (CBD) fused at the C-terminus of the intein domain to facilitate purification. As shown in Fig. 2A, both intein precursors had comparable levels of expression in *E. coli* cells ($\sim 5 \text{ mg L}^{-1}$ as estimated by UV spectroscopy), and showed similar rates and propensities for *in vivo* cleavage ($\sim 65\%$). It is worth noting that under these conditions most of the intein fusion precursors in both cases were expressed as soluble proteins (Fig. 2A).

We next tested the ability of the different precursors to be cleaved *in vitro* by using reduced glutathione (GSH) to produce folded RTD-1. GSH has been shown to promote cyclization and concomitant folding when used in the biosynthesis of Cys-rich cyclic polypeptides.^{21,25,26} The cyclization/folding reaction was performed on the chitin beads where the corresponding precursors had been purified. The best cleavage/cyclization conditions were accomplished using 100 mM of GSH in phosphate buffer at pH 7.2 for 48 h at room temperature. Under these conditions both RTD-intein precursors were completely cleaved (Fig. 2A). HPLC analysis of the crude cyclization mixture revealed that in both cases the main peptide product was the corresponding folded RTD-1 (Fig. 2B) as revealed by HPLC and ES-MS analysis (Fig. 2C and D). Other peptide minor peaks in the HPLC chromatograms were identified as not correctly folded GSH-adducts. Precursor RTD-C3 gave the best cyclization/folding yield producing around $10 \mu\text{g L}^{-1}$ of purified RTD-1. This yield corresponds to $\sim 10\%$ of the theoretical value. The *in vitro* cyclization/folding of precursor RTD-C7 gave a slightly lower yield ($\sim 7 \mu\text{g L}^{-1}$ of purified RTD-1). The lower yield observed for this precursor can be attributed to the slightly more complex cyclization/folding crude (Fig. 2B).

Recombinant RTD-1 was purified by HPLC and its biological activity tested using an anthrax LF inhibitory assay.²⁸ *In vitro* produced RTD-1 was able to inhibit anthrax LF with an IC_{50} of $384 \pm 33 \text{ nM}$ (Fig. 3A), which corresponds to a $K_i \approx 0.4 \mu\text{M}$ under the conditions used in the inhibitory assay.^{29,30} This IC_{50} value closely parallels previously reported values,^{20,31} thus confirming the biological activity of the recombinantly produced θ -defensin.

We also used NMR spectroscopy to confirm that recombinant RTD-1 adopted a native θ -defensin fold. The structure of native

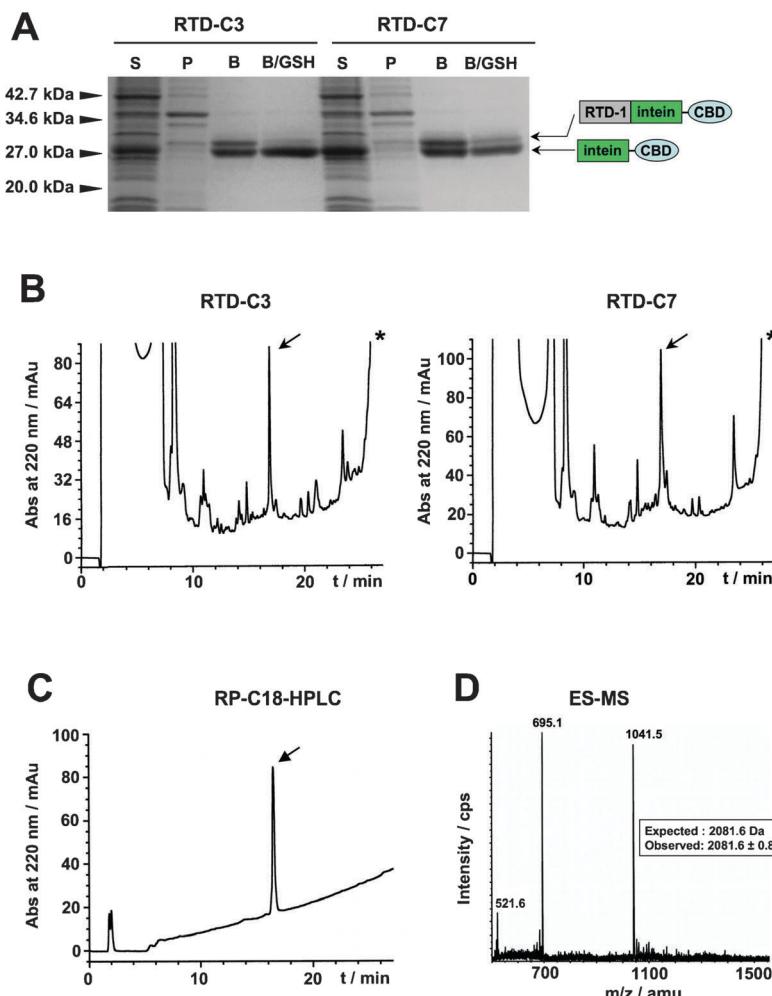


Fig. 2 *In vitro* production of RTD-1. (A) SDS-PAGE analysis of cell lysates from BL21 *E. coli* cells expressing precursors RTD-C3 and RTD-C7. The identity of the bands corresponding to intein precursors and *in vivo* cleaved proteins is shown on the right. The positions and molecular weights of the molecular markers used are shown on the left of the gel (P: insoluble cell lysate, S: soluble cell lysate, B: purified intein, B/GSH: purified intein after cleavage with 100 mM glutathione at pH 7.2). (B) Reverse-phase C18-HPLC traces for the GSH-induced cyclization folding of purified precursors RTD-C3 and RTD-C7. The peak corresponding to the folded RTD-1 is indicated with an arrow in each case. The asterisk denotes the cleaved intein-CBD fragment. (C) Reverse-phase C18-HPLC of purified RTD-1. (D) ES-MS spectrum of purified RTD-1. Calculated mass corresponds to the average isotopic mass.

RTD-1 has been previously reported by Craik *et al.* in 10% MeCN aqueous buffer at pH 4.5.⁶ We decided to carry out the NMR experiments under more physiological conditions using an aqueous buffer containing no organic solvents at pH 6.5. The rationale to use more physiological conditions was to allow the future study of biologically relevant interactions between RTD-1 and potential biomolecular targets by NMR. The chemical shifts of the assigned backbone amide and alpha protons (HN α and HC α) for recombinant RTD-1 at pH 6.5 were very similar to those reported earlier by Craik at pH 4.5 (Table S2, ESI \dagger). No significant (≥ 0.3 ppm) differences were found in the backbone amide protons. We also saw a uniform shift rather than variable changes for the backbone alpha protons (~ 0.2 ppm) at pH 6.0 (Table S2, ESI \dagger). Since backbone alpha protons usually reflect the secondary structure of the peptide backbone, this uniform offset in the resonances of the backbone HC α protons could be attributed to the different buffer conditions used in the two samples rather than changes in secondary structure.

We also produced recombinant ^{15}N -labeled RTD-1 by expressing the RTD-C3 precursor in minimal M9 medium containing $^{15}\text{NH}_4\text{Cl}$ as the only source of nitrogen. Under these conditions the expression yield was around $7 \mu\text{g L}^{-1}$ of ^{15}N -labeled RTD-1 after purification by HPLC. The HSQC spectrum of recombinant RTD-1 was very well dispersed, indicating a well-folded structure (Fig. 3B). Having access to the recombinant expression of θ -defensins allows the introduction of NMR active isotopes (^{15}N and/or ^{13}C) in a very inexpensive fashion, thus facilitating the use of heteronuclear NMR to study intermolecular interactions between θ -defensins and their biomolecular targets.

Encouraged by the results obtained with the *in vitro* GSH-induced cyclization/folding of the RTD-intein precursors we also decided to explore the expression of folded RTD-1 inside *E. coli* cells. RTD-1 has been shown to be antimicrobial against both Gram-positive and Gram-negative bacteria. However, the antimicrobial activity of RTD-1 has been shown to strongly depend on the conditions used in the antimicrobial assays. The presence

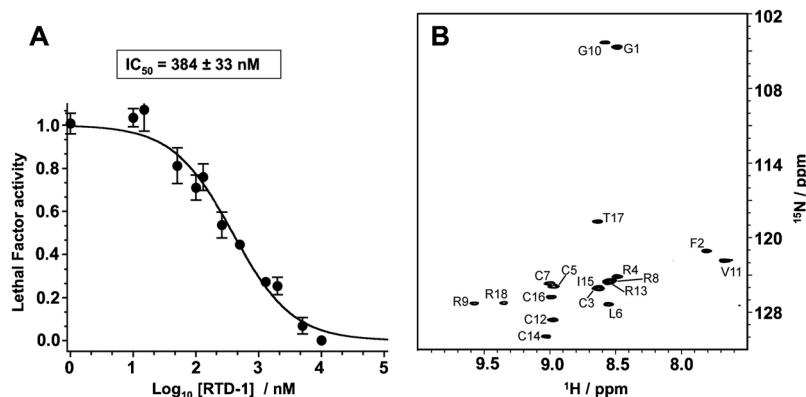


Fig. 3 Characterization of recombinantly produced θ -defensin RTD-1. (A) Inhibition assay of RTD-1 against anthrax LF. Different concentrations of RTD-1 were tested against LF. At each concentration, residual LF activity was measured and divided by the activity of LF in the absence of inhibitor. Activity was measured as the rate LF protease cleaves a fluorescence LF substrate²⁸ and determined by the rate of fluorescence signal decaying (see experimental). (B) ¹⁵N-HSQC spectrum of ¹⁵N-labeled RTD-1 defensin in water at pH 6.0. The identity of the crosspeaks is indicated by the number of the residue according to Fig. 1.

of 10% human serum has been shown to significantly decrease the antimicrobial properties of RTD-1 especially against Gram-negative bacteria such as *E. coli*.¹⁴ We anticipated that the high molecular complexity of the bacterial cytosol could decrease the antimicrobial activity of RTD-1 when produced intracellularly and therefore allow its production in the cellular cytosol.

In order to test this hypothesis we used the RTD-C7 precursor (Fig. 1B). In-cell expression of RTD-1 was accomplished in Origami2(DE3). These cells have mutations in the thioredoxin and glutathione reductase genes, which facilitates the formation of disulfide bonds in the bacterial cytosol.³² We have recently used these cells for the *in vivo* production of several disulfide-containing backbone cyclized polypeptides.^{22,25,26}

Precursor RTD-C7 was expressed in Origami2(DE3) overnight at room temperature giving a total yield of precursor protein of ~ 3 mg L⁻¹. Under these conditions the precursor was completely cleaved *in vivo* (Fig. 4A). In contrast when precursor RTD-C3 was expressed under these conditions only $\sim 74\%$ of the precursor protein was cleaved *in vivo* (Fig. 4A). This difference could be attributed to the proximity of Cys⁵ to the RTD-intein junction in the precursor RTD-C7, which is only one residue away from the RTD-intein junction therefore facilitating the intramolecular cleavage. In the precursor RTD-C3, the closest Cys residue within the sequence (Cys¹⁶) is four residues away from the RTD-intein junction (Fig. 1). After lysing the cells expressing RTD-C7, both the insoluble and soluble cellular fractions were analyzed by HPLC-MS/MS in multiple reaction mode to identify and quantify the amount of folded RTD-1. Interestingly, almost all of the folded RTD-1 produced inside living *E. coli* cells was found in the insoluble fraction (~ 22 μ g L⁻¹, Fig. 4B). θ -Defensins are known to interact with model phospholipid membranes, and have been shown to have low nM affinity for glycoproteins (gp120 and cd4) and glycolipids (galactosylceramide).¹⁹ This could explain the higher affinity of recombinant RTD-1 for the insoluble cellular lysate fraction, where it is likely to find insoluble membrane fragments containing glycolipids and/or membrane-bound peptidoglycan precursors. In fact, peptidoglycan precursor and membrane-anchored lipid II have been shown to bind to human α - and β -defensins.³³

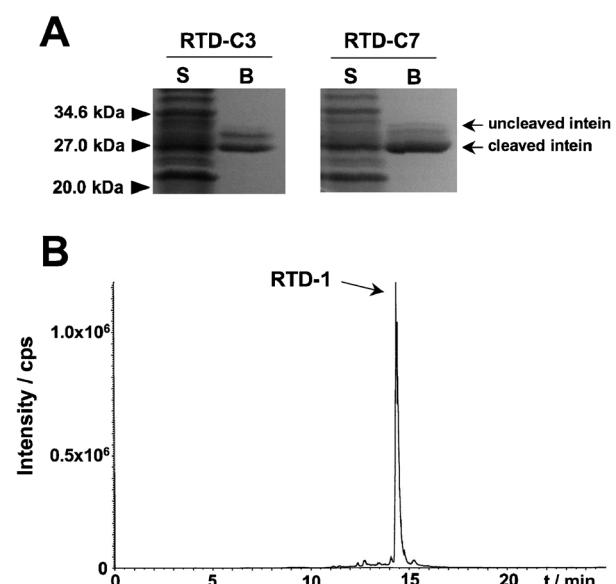


Fig. 4 *In vivo* production of RTD-1 in *E. coli* cells. (A) SDS-PAGE analysis of the soluble cell lysate fraction (S) of Origami *E. coli* cells expressing precursors RTD-C3 and RTD-C7. Expression was induced as described in the text at room temperature for 18 h (S: soluble cell lysate and B: purified intein). (B) HPLC-MS/MS analysis in multiple reaction mode of *in vivo* produced RTD-1 in the insoluble fraction of the cell lysate. Specific MRM for RTD-1 identified ~ 22 μ g L⁻¹ of folded RTD-1.

In summary, we report in this work the first recombinant expression of the θ -defensin RTD-1 in *E. coli* cells. RTD-1 can be produced either *in vitro*, by processing of the corresponding RTD-intein precursor, or *in vivo* by using longer expression times in combination with Origami or similar *E. coli* cells. The yield of purified RTD-1 was greater *in vivo* than *in vitro* (~ 22 μ g vs. ~ 10 μ g of purified peptide per L of bacterial culture). Although the yields are modest, our approach represents an improvement over the approach reported by Suga *et al.* for the ribosomal synthesis of RTD-1 using a cell-free expression system in combination with genetic code reprogramming.³¹ The ribosomal

synthesis of RTD-1 using cell-free expression systems usually provides low yields and is expensive to scale-up. Although this approach allows the production of libraries, these libraries cannot be easily decoded requiring time consuming deconvolution methods.³¹

The chemical synthesis of θ -defensin RTD-1 has also been accomplished by either Fmoc-¹ or Boc-based⁶ solid-phase peptide synthesis (SPPS). Both approaches, however, are rather complex and expensive to scale-up. Because of the high cost associated with chemical synthesis and relatively poor yields of the cell-free expression system, bacterial expression provides an ideal solution for its large-scale economic production. Bacterial expression of RTD-1 could be easily scaled-up by using richer media and/or fermentors to allow expression at higher cellular densities. The bacterial expression of RTD-1 also allows the incorporation of NMR-active isotopes in a very inexpensive manner to facilitate the structural study of the interactions between θ -defensin and potential biomolecular targets through heteronuclear NMR. For example, we are using ¹⁵N-labeled RTD-1 to study the potential interaction of this defensin with a soluble form of lipid II. More interestingly, in-cell expression of RTD-1 gave a yield of $\sim 22 \mu\text{g L}^{-1}$, which corresponds to an intracellular concentration of $\sim 4 \mu\text{M}$. This result opens the intriguing possibility for the generation of genetically-encoded libraries using the θ -defensin scaffold. These libraries could be used for the development of in-cell screening and directed evolution technologies to further study the biological activity of θ -defensins or to engineer new biological activities on the defensin scaffold.

Experimental

Analytical reverse phase (RP)-HPLC was performed on a HP1100 series instrument with 220 nm and 280 nm detection using a Vydac C18 column (5 mm, 4.6 \times 150 mm) at a flow rate of 1 mL min⁻¹. Semipreparative RP-HPLC was performed on a Waters Delta Prep system fitted with a Waters 2487 Ultraviolet-Visible (UV-vis) detector using a Vydac C18 column (5 μm , 10 \times 250 mm) at a flow rate of 5 mL min⁻¹. All runs used linear gradients of 0.1% aqueous trifluoroacetic acid (TFA, solvent A) vs. 0.1% TFA, 90% MeCN in H₂O (solvent B). UV-vis spectroscopy was carried out on an Agilent 8453 diode array spectrophotometer. Electrospray mass spectrometry (ES-MS) was performed on an Applied Biosystems API 3000 triple quadrupole mass spectrometer using Analyst 1.4.2. Calculated masses were obtained using Analyst 1.4.2. Protein samples were analyzed by SDS-PAGE 4–20% Tris-Glycine Gels (Lonza, Rockland, ME). The gels were then stained with Pierce (Rockford, IL) Gelcode Blue, photographed/digitized using a Kodak (Rochester, NY) EDAS 290, and quantified using NIH Image-J software (<http://rsb.info.nih.gov/ij/>). The integrity of all plasmids was confirmed by DNA sequencing. DNA sequencing was performed by the DNA Sequencing and Genetic Analysis Core Facility at the University of Southern California using an ABI 3730 DNA sequencer, and the sequence data were analyzed with DNASTar (Madison, WI) Lasergene v8.0.2. Amino acid analysis was performed at the Amino Acid Laboratory in the Department of Molecular Biosciences, School of Veterinary Medicine,

University of California at Davis. All chemicals were obtained from Sigma-Aldrich (Milwaukee, WI) unless otherwise indicated.

Cloning of RTD-C3 and RTD-C7 and *in vitro* production of RTD-1

Synthetic DNA oligos (Integrated DNA technologies, Coralville, IA) encoding RTD-1 (Table S1, ESI†) was annealed and ligated into the pTXB1 vector (New England Biolabs, Ipswich, MA) using the NdeI and SapI restriction sites as described previously.^{21,22} The resulting plasmids were transformed into either BL21(DE3) or Origami2(DE3) cells (EMD Chemicals, Gibbstown, NJ) and grown in LB broth. Transformed BL21(DE3) cells were induced with 0.3 mM IPTG for 3 h at 30 °C and transformed Origami2(DE3) cells with 0.1 mM IPTG for 20 h at 22 °C. Cells were lysed in 0.1 mM EDTA, 1 mM PMSF, 50 mM sodium phosphate, 250 mM sodium chloride buffer at pH 7.2 containing 5% glycerol by sonication. The soluble fraction was incubated with chitin beads (New England Biolabs) for 1 h at 4 °C and the beads were washed with column buffer (0.1 mM EDTA, 50 mM sodium phosphate, 250 mM sodium chloride buffer at pH 7.2) containing 0.1% Triton X-100 followed by washes with column buffer without Triton X-100. The peptide was cyclized and folded *in vitro* using column buffer at pH 7.2 containing 100 mM reduced glutathione (GSH) for 2–3 days at room temperature with gentle rocking. We found that under these conditions RTD-1 binds strongly to the chitin column, and therefore was eluted using 8 M GdmCl in water. The corresponding supernatant and washes were pooled, and RTD-1 was purified by semipreparative HPLC using a linear gradient of 17–39% solvent B for over 30 min. Purified products were characterized by HPLC and ES-MS. RTD-1 was quantified by amino acid analysis. The expression of RTD-intein fusion precursors was quantified first by desorption of the proteins from an aliquot of chitin beads using 8 M GdmCl and then measurement by UV-VIS using a molar absorption coefficient value of 36 660 M⁻¹ cm⁻¹.

Expression of ¹⁵N-labeled RTD-1

Expression was carried out using BL21(DE3) cells as described above except those grown in M9 minimal medium containing 0.1% ¹⁵NH₄Cl as the nitrogen source.^{25,34} Cyclization and folding were performed as described above. The ¹⁵N-labeled RTD-1 was purified by semipreparative HPLC as before. Purified products were characterized by HPLC and ES-MS (Fig. S1, ESI†).

In-cell expression of RTD-1

Origami2(DE3) cells transformed with the plasmid encoding the RTD-C7 precursor were grown, induced, harvested and lysed as described above. The insoluble pellet was first washed three times with column buffer containing 0.1% Triton X-100 and then twice with just column buffer. The resulting pellet was dissolved in minimal amount of 8 M GdmCl in water. Both the soluble cell lysate and solubilized cell lysate pellet were extracted using C18 SepPak cartridges (Waters, Milford, MA) with elution in MeCN:H₂O (3:2 vol) containing 0.1% TFA and lyophilized. The samples were dissolved in H₂O containing 0.1% formic acid and analyzed by HPLC-tandem mass spectrometry using a C18-HPLC column (5 mm, 2.1 \times 100 mm),

and $\text{H}_2\text{O}-\text{MeCN}$ buffers containing 0.1% formic acid as mobile phase. Typical analysis used a linear gradient of 0–90% MeCN in H_2O for over 10 min. Detection was performed on an API 3000 ES-MS using a multiple reaction-monitoring (MRM) mode. MRM analysis was performed using peaks at $m/z = 521.6$ (4th charge state) and 695.1 (3rd charge state). Data were collected and processed using Analyst software (Applied Biosystems). The calibration curve using pure RTD-1 was found to be linear in the range of 10–40 ng. Loss of RTD-1 during extraction was quantified by spiking a control sample with a known amount of purified RTD-1 and analysis by HPLC-MS/MS. The recovery was found to be approximately 45%. Intracellular RTD-1 concentration was calculated using the following constants for *E. coli* K12: intracellular volume $0.58 \times 10^{-15} \text{ L cell}^{-1}$; wet weight $640 \times 10^{-15} \text{ g cell}^{-1}$ (<http://bionumbers.hms.harvard.edu>).

NMR spectroscopy

NMR samples were prepared by dissolving ^{15}N -labeled backbone cyclized RTD-1 into 80 mM potassium phosphate buffer at pH 6.0 in 90% $\text{H}_2\text{O}/10\% \text{D}_2\text{O}$ (v/v) or 100% D_2O to a concentration of approximately 0.2 mM. All ^1H NMR data were recorded on a Bruker Avance II 500 MHz spectrometer equipped with a cryoprobe. Data were acquired at 25 °C, and 2,2-dimethyl-2-silapentane-5-sulfonate, DSS, was used as an internal reference. All 2D $^1\text{H}\{^1\text{H}\}$ -TOCSY and $^1\text{H}\{^1\text{H}\}$ -NOESY and 3D experiments, $^1\text{H}\{^{15}\text{N}\}$ -TOCSY-HSQC and $^1\text{H}\{^{15}\text{N}\}$ -NOESY-HSQC, were performed according to standard procedures³⁵ with spectral widths of 14 ppm in proton dimensions and 35 ppm in nitrogen dimension. The carrier frequency was centered on the water signal, and the solvent was suppressed by using WATERGATE pulse sequence. TOCSY (spin lock time 80 ms) and 2D-NOESY (mixing time 250 ms) spectra were collected using 4096 t_2 points and 512 t_1 blocks of 64 transients. 3D-TOCSY-HSQC (spin lock time 80 ms) and 3D-NOESY-HSQC (mixing time 250 ms) spectra were collected using 1024 t_3 points, 256 t_2 and 128 t_1 blocks of 16 transients. Spectra were processed using Topspin 1.3 (Bruker). Each 3D data set was apodized by 90°-shifted sinebell-squared in all dimensions, and zero filled to $1024 \times 512 \times 256$ points prior to Fourier transformation. Assignments for the backbone nitrogens, H^α and H' protons (Tables S2, ESI†) were obtained using standard procedures.^{35,36}

Anthrax lethal factor protease inhibition assay

Lethal factor (LF) protease and FRET-based substrate containing fluorescent proteins CyPet and YPet linked by consensus sequence (RRKKVYPYPMEGTIA) were expressed and purified as previously described.^{28,37} RTD-1 concentrations were quantified by amino acid analysis. LF inhibition assay was performed in LF reaction buffer (10 μM CaCl_2 , 10 μM MgCl_2 , 20 μM ZnCl_2 , 20 mM sodium phosphate, and 100 mM NaCl at pH 7.2). Samples of 50 nM LF in LF reaction buffer (100 μL) were preincubated with different concentrations of RTD-1 ranging from 1 nM to 10 μM . After incubation at room temperature for 30 min, residual LF activity was measured by adding the FRET-based substrate to a final concentration of 10 nM and the decrease in FRET signal was observed.²⁸ Measurements of the FRET signal

were taken every 2 min for 3 h. FRET was measured using an Envision 2103 plate reader (PerkinElmer) using an excitation wavelength of 405 nm. The relative FRET change was calculated using: $\text{FRC} = I_t^{535}/I_0^{535}$, where I_0 and I_t are the fluorescence intensities at time zero and at a particular time (t), at 535 nm. The initial velocities for the hydrolysis of substrate lethal factor in the presence of different concentrations of RTD-1 were fitted to a one-site competitive binding equation using the software package Prism (GraphPad Software). K_i was calculated using the equation of Cheng and Prusoff²⁹ and a K_m value of 40 μM .³⁰

Acknowledgements

This work was supported by National Institutes of Health Research Grant R01-GM090323 (JAC) and NIH Grant 5R01GM085006-02 (AS); and by the Department of Defense Congressionally Directed Medical Research Program Grant PC09305 (JAC).

Notes and references

- Y. Q. Tang, J. Yuan, G. Osapay, K. Osapay, D. Tran, C. J. Miller, A. J. Ouellette and M. E. Selsted, *Science*, 1999, **286**, 498–502.
- H. Tanabe, J. Yuan, M. M. Zaragoza, S. Dandekar, A. Henschel-Edman, M. E. Selsted and A. J. Ouellette, *Infect. Immun.*, 2004, **72**, 1470–1478.
- M. E. Selsted and A. J. Ouellette, *Nat. Immunol.*, 2005, **6**, 551–557.
- R. I. Lehrer, *Nat. Rev. Microbiol.*, 2004, **2**, 727–738.
- K. J. Rosengren, N. L. Daly, L. M. Fornander, L. M. Jonsson, Y. Shirafuji, X. Qu, H. J. Vogel, A. J. Ouellette and D. J. Craik, *J. Biol. Chem.*, 2006, **281**, 28068–28078.
- M. Trabi, H. J. Schirra and D. J. Craik, *Biochemistry*, 2001, **40**, 4211–4221.
- L. Leonova, V. N. Kokryakov, G. Aleshina, T. Hong, T. Nguyen, C. Zhao, A. J. Waring and R. I. Lehrer, *J. Leukocyte Biol.*, 2001, **70**, 461–464.
- D. Tran, P. A. Tran, Y. Q. Tang, J. Yuan, T. Cole and M. E. Selsted, *J. Biol. Chem.*, 2002, **277**, 3079–3084.
- P. Tongaonkar, P. Tran, K. Roberts, J. Schaal, G. Osapay, D. Tran, A. J. Ouellette and M. E. Selsted, *J. Leukocyte Biol.*, 2011, **89**, 283–290.
- T. X. Nguyen, A. M. Cole and R. I. Lehrer, *Peptides*, 2003, **24**, 1647–1654.
- A. E. Garcia, G. Osapay, P. A. Tran, J. Yuan and M. E. Selsted, *Infect. Immun.*, 2008, **76**, 5883–5891.
- C. Stegemann, E. V. Tsvetkova, G. M. Aleshina, R. I. Lehrer, V. N. Kokryakov and R. Hoffmann, *Rapid Commun. Mass Spectrom.*, 2010, **24**, 599–604.
- A. M. Cole, T. Hong, L. M. Boo, T. Nguyen, C. Zhao, G. Bristol, J. A. Zack, A. J. Waring, O. O. Yang and R. I. Lehrer, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 1813–1818.
- D. Tran, P. Tran, K. Roberts, G. Osapay, J. Schaal, A. Ouellette and M. E. Selsted, *Antimicrob. Agents Chemother.*, 2008, **52**, 944–953.
- W. T. Penberthy, S. Chari, A. L. Cole and A. M. Cole, *Cell. Mol. Life Sci.*, 2011, **68**, 2231–2242.
- C. Munk, G. Wei, O. O. Yang, A. J. Waring, W. Wang, T. Hong, R. I. Lehrer, N. R. Landau and A. M. Cole, *AIDS Res. Hum. Retroviruses*, 2003, **19**, 875–881.
- S. M. Owen, D. Rudolph, W. Wang, A. M. Cole, M. A. Sherman, A. J. Waring, R. I. Lehrer and R. B. Lal, *J. Pept. Res.*, 2004, **63**, 469–476.
- S. A. Gallo, W. Wang, S. S. Rawat, G. Jung, A. J. Waring, A. M. Cole, H. Lu, X. Yan, N. L. Daly, D. J. Craik, S. Jiang, R. I. Lehrer and R. Blumenthal, *J. Biol. Chem.*, 2006, **281**, 18787–18792.
- W. Wang, A. M. Cole, T. Hong, A. J. Waring and R. I. Lehrer, *J. Immunol.*, 2003, **170**, 4708–4716.
- W. Wang, C. Mulakala, S. C. Ward, G. Jung, H. Luong, D. Pham, A. J. Waring, Y. Kaznessis, W. Lu, K. A. Bradley and R. I. Lehrer, *J. Biol. Chem.*, 2006, **281**, 32755–32764.

21 R. H. Kimura, A. T. Tran and J. A. Camarero, *Angew. Chem., Int. Ed.*, 2006, **45**, 973–976.

22 J. A. Camarero, R. H. Kimura, Y. H. Woo, A. Shekhtman and J. Cantor, *ChemBioChem*, 2007, **8**, 1363–1366.

23 J. A. Camarero and T. W. Muir, *J. Am. Chem. Soc.*, 1999, **121**, 5597–5598.

24 J. A. Camarero, D. Fushman, D. Cowburn and T. W. Muir, *Bioorg. Med. Chem.*, 2001, **9**, 2479–2484.

25 J. Austin, W. Wang, S. Puttamadappa, A. Shekhtman and J. A. Camarero, *ChemBioChem*, 2009, **10**, 2663–2670.

26 J. Austin, R. H. Kimura, Y. H. Woo and J. A. Camarero, *Amino Acids*, 2010, **38**, 1313–1322.

27 J. A. Camarero, J. Pavel and T. W. Muir, *Angew. Chem., Int. Ed.*, 1998, **37**, 347–349.

28 R. H. Kimura, E. R. Steenblock and J. A. Camarero, *Anal. Biochem.*, 2007, **369**, 60–70.

29 Y. Cheng and W. H. Prusoff, *Biochem. Pharmacol.*, 1973, **22**, 3099–3108.

30 F. Tonello, P. Ascenzi and C. Montecucco, *J. Biol. Chem.*, 2003, **278**, 40075–40078.

31 T. Kawakami, A. Ohta, M. Ohuchi, H. Ashigai, H. Murakami and H. Suga, *Nat. Chem. Biol.*, 2009, **5**, 888–890.

32 P. H. Bessette, F. Aslund, J. Beckwith and G. Georgiou, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 13703–13708.

33 M. Wilmes, B. P. Cammue, H. G. Sahl and K. Thevissen, *Nat. Prod. Rep.*, 2011, **28**, 1350–1358.

34 S. S. Puttamadappa, K. Jagadish, A. Shekhtman and J. A. Camarero, *Angew. Chem., Int. Ed.*, 2010, **49**, 7030–7034.

35 J. Cavanagh and M. Rance, *J. Magn. Reson.*, 1992, **96**, 670–678.

36 K. Wuethrich, *NMR of Proteins and Nucleic Acids*, 1986.

37 Y. Kwon, K. Welsh, A. R. Mitchell and J. A. Camarero, *Org. Lett.*, 2004, **6**, 3801–3804.



Efficient one-pot cyclization/folding of Rhesus θ -defensin-1 (RTD-1)

Teshome L. Aboye ^{a,†}, Yilong Li ^{a,†}, Subhabrata Majumder ^c, Jinfeng Hao ^a, Alexander Shekhtman ^c, Julio A. Camarero ^{a,b,*}

^a Department of Pharmacology and Pharmaceutical Sciences, University of Southern California, 1985 Zonal Avenue, PSC 616, Los Angeles, CA 90033, USA

^b Department of Chemistry, University of Southern California, 1985 Zonal Avenue, PSC 616, Los Angeles, CA 90033, USA

^c Department of Chemistry, State University of New York, Albany, NY 12222, USA

ARTICLE INFO

Article history:

Received 3 January 2012

Revised 18 February 2012

Accepted 23 February 2012

Available online 2 March 2012

Keywords:

Defensin

Cyclic peptide

Antimicrobial peptide

ABSTRACT

We report an efficient approach for the chemical synthesis of Rhesus θ -defensin-1 (RTD-1) using Fmoc-based solid-phase peptide synthesis in combination with an intramolecular version of native chemical ligation. The corresponding linear thioester precursor was cyclized and folded in a one-pot reaction using reduced glutathione. The reaction was extremely efficient yielding natively folded RTD-1 with minimal or no purification at all. This approach is fully compatible with the high throughput production of chemical libraries using this peptide scaffold.

© 2012 Elsevier Ltd. All rights reserved.

Defensins are antimicrobial peptides that are important in the innate immune defense of mammals.^{1–3} Mammalian defensins have broad-spectrum antimicrobial activities including anti-viral, and have been shown to be involved in other defense mechanisms, such as immune modulation, neutralization of endotoxin, and anti-cancer activities, among others.^{3,4} Mammalian defensins are cationic cysteine-rich peptides with largely β -sheet structures stabilized by the presence of three disulfide bonds, and can be classified into three structurally distinct groups, α -, β - and θ -defensins. Despite differences in disulfide connectivities and the presence of an N-terminal α -helix segment in β -defensins, α - and β -defensins adopt a similar overall fold. θ -Defensins, on the other hand, are backbone-cyclized octadapeptides that are produced by the pair-wise, head-to-tail splicing of nonapeptides derived from α -defensin related precursors by an as yet unknown mechanism.^{1,5} θ -Defensins are only expressed in Old World monkeys and orangutans,^{6–8} and to date, they remain the only known backbone-cyclized polypeptides expressed in animals.¹ Interestingly, humans possess genes encoding θ -defensins, but they have lost the ability to produce the peptides due to a stop codon mutation within the signal sequence that prevents subsequent translation.⁹

Rhesus θ -defensin-1 (RTD-1) was the first θ -defensin to be found in leukocyte extracts from Rhesus macaques.¹ Other less abundant RTD variants, named RTD-2 through RTD-6, have also been found in Rhesus macaques more recently.^{10–12} θ -Defensins have broad antimicrobial activities in the presence of physiological

concentrations of salt, divalent cations, and serum *in vitro*.^{1,13} Interestingly, they are also anti-fungal¹ and have anti-viral activities against herpes simplex¹⁴ and HIV.^{9,15} In fact, chemically-synthesized θ -defensins derived from the corresponding human pseudogene sequences have been shown to protect human cells from infection by HIV-1⁹ and have been evaluated as a topical anti-HIV agent for the prevention of HIV transmission.^{16–18} θ -Defensins also inactivate germinating anthrax spores and act as a competitive inhibitor of anthrax lethal factor protease.¹⁹

The structure of RTD-1 (Fig. 1A) is different from that of α - and β -defensins, exhibiting a backbone-cyclized double β -hairpin-like fold with two anti-parallel β -strands stabilized by three disulfide bonds in a ladder configuration (Fig. 1A).²⁰ The circular backbone topology of θ -defensins is required for their biological activity as their linear counterparts show a decreased biological activity. The presence of three disulfides and a circular backbone topology makes θ -defensins a particularly stable peptide framework. Altogether, these features make these peptides an attractive molecular scaffold for the development of peptide-based therapeutics with optimized antimicrobial activity or for the introduction of novel biological activities.

The relatively small size of θ -defensins (18 residues) makes it possible to use chemical tools for the generation of large combinatorial libraries based on this scaffold for the screening and selection of optimized sequences for a particular biological activity. The chemical synthesis of θ -defensin RTD-1 has already been accomplished by either Fmoc-¹ or Boc-based²⁰ solid-phase peptide synthesis (SPPS). In both approaches, the fully reduced unprotected linear peptide was first folded under physiological conditions and then the N-terminal amino and C-terminal carboxylic groups were

* Corresponding author. Tel.: +1 323 442 1417.

E-mail address: jcamarer@usc.edu (J.A. Camarero).

† Authors contributed equally to this work.

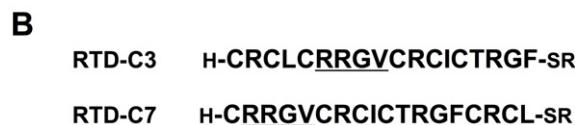
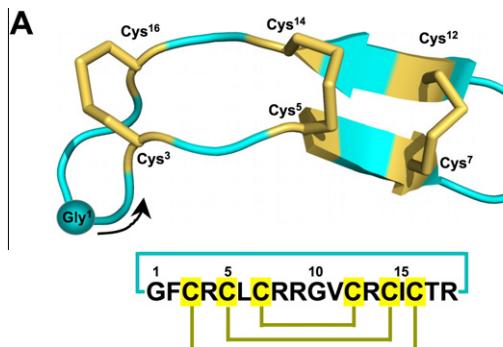
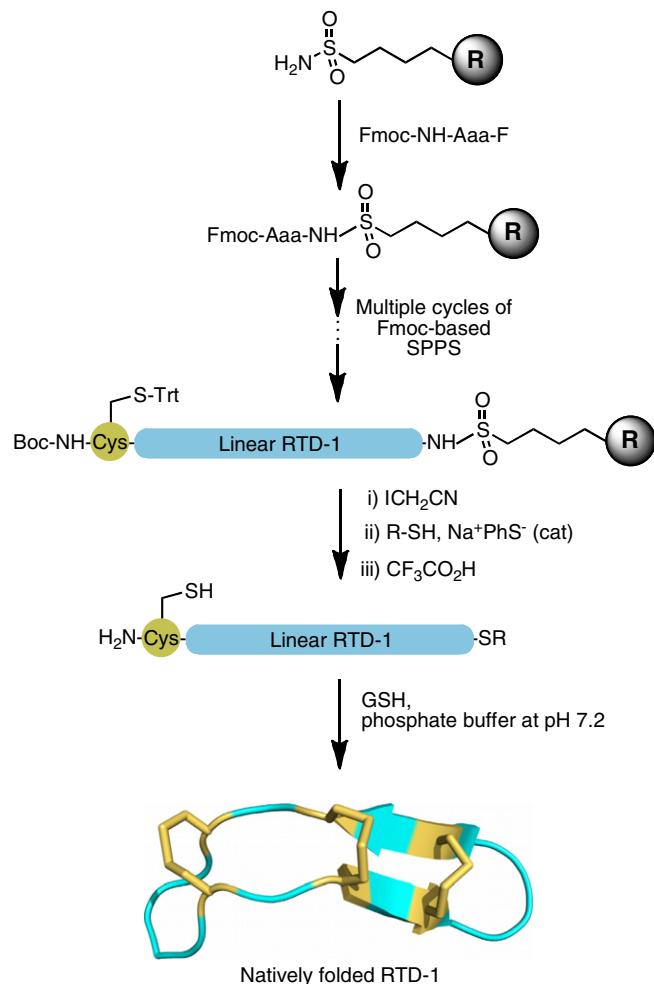


Figure 1. (A) Primary and tertiary structures of rhesus θ -defensin 1 (RTD-1) (PDB ID code: 1HVZ).²⁰ The backbone cyclized peptide (connecting bond shown in blue) is stabilized by the three-disulfide bonds in a ladder formation (disulfide bonds shown in yellow). A solid circle is used to indicate the position of residue Gly¹ with an arrow indicating the N-to-C direction of the polypeptide backbone. (B) Sequences of the linear peptide thioester precursors used in this work, RTD-C3 and RTD-C7. The sequence RRGV is underlined for reference.

reacted using different coupling reagents.^{1,20} However, due to the use of non-chemoselective conditions during the backbone cyclization of the un-protected peptide, the ligation reaction required careful optimization and in both cases the final product had to be purified. Although this approach has been successful in the chemical production of different θ -defensins, the production of large libraries would require the optimization and purification of every potential member of the library making this approach impractical. In order to overcome this problem we decided to explore the use of intramolecular native chemical ligation (NCL)^{21,22} for the efficient production of cyclic θ -defensins (Scheme 1). Intramolecular NCL requires the presence of an N-terminal Cys residue and a C-terminal α -thioester group in the same linear precursor molecule.^{23,24} The chemoselectivity of this intramolecular reaction is extremely exquisite at neutral pH producing the backbone cyclized peptide in almost quantitative yields.^{24–26}

RTD-1 has six Cys residues that may be used for cyclization. In order to facilitate the cyclization reaction, we decided to use the Cys residues located in positions 3 and 7 (Fig. 1). These Cys residues do not possess a charged or β -branched residue N-terminally adjacent, which should facilitate the kinetics of the cyclization reaction. Accordingly, two different RTD-1 linear thioester precursors (RTD-C3 and RTD-C7) were produced by Fmoc-based SPPS using a safety-catch sulfonamide linker (Scheme 1 and Fig. 1B).^{27,28} Activation of the sulfonamide peptide-resins with iodoacetonitrile followed by thiolytic cleavage with ethyl 3-mercaptopropionate provided the fully protected C-terminal thioester peptide precursors. Subsequent acidolysis with trifluoroacetic acid (TFA), and purification by precipitation with Et₂O gave the corresponding unprotected linear thioester precursors with yields \approx 10–15% based on the initial substitution of the resin. As shown in Figure 2A, HPLC analysis of the TFA crude for both linear precursors revealed in both cases the presence of a major peak that corresponded to the expected linear thioester precursor as determined by electro-spray mass spectrometry (ESMS) analysis (Fig. 2A). Other minor peaks present in both TFA crudes were identified by ES-MS as partially oxidized linear precursors.

We next performed the backbone cyclization/folding in a one-pot reaction by diluting the corresponding unpurified peptide thioester precursors to a final peptide concentration of \approx 25 μ M



Scheme 1. Diagram depicting the chemical synthesis of RTD-1 by Fmoc-based SPPS followed by GSH-induced cyclization/folding.

in sodium phosphate buffer at pH 7.0 in the presence of 1 mM reduced glutathione (GSH). GSH has been shown to promote cyclization and concomitant folding when used in the production of several Cys-rich cyclic polypeptides.^{26,29–31} As shown in Figure 2A, the GSH-induced cyclization/folding reaction was very clean and efficient for both linear precursors, being complete in 24 h. In both cases the major peptide product was identified by ES-MS as the expected folded cyclized RTD-1 product (Fig. 2). It is worth noting, that the cyclization/folding of precursor RTD-C7 was extremely efficient producing a product with more than 90% of the peptide content being the desired cyclic folded product therefore requiring no purification (Fig. 2A). The yield for the one-pot cyclization and folding using precursor RTD-C3 was estimated by HPLC to be \approx 80% (Fig. 2B).

The biological activity of pure synthetic RTD-1 was tested using an anthrax lethal factor (LF) protease inhibitory assay.³² Synthetic RTD-1 was able to inhibit LF with an IC₅₀ of 419 \pm 32 nM (Fig. 3A). This IC₅₀ value corresponds to a $K_i \approx 0.4 \mu$ M under the conditions used in the inhibitory assay.^{33,34} This IC₅₀ is similar to the values previously reported for native RTD-1,^{19,35} therefore confirming the biological activity of synthetic RTD-1.

Pure synthetic RTD-1 was also characterized by 2D-¹H NMR to confirm its θ -defensin fold. The structure of RTD-1 has been previously elucidated by NMR in 10% MeCN aqueous buffer at pH 4.5.²⁰ We have also recently characterized native folded recombinant ¹⁵N-labeled RTD-1 in phosphate buffer at pH 6.0.³⁶ The chemical

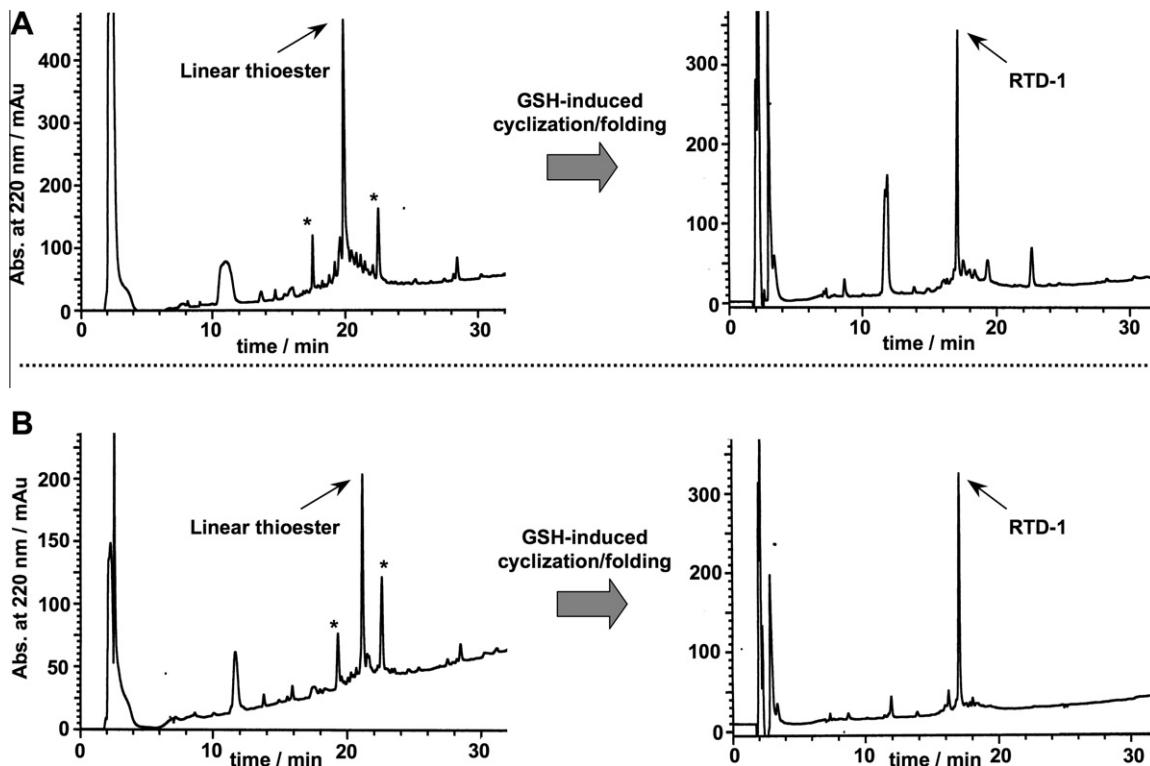


Figure 2. Reverse-phase C18-HPLC traces for the GSH-induced cyclization folding of purified precursors RTD-C3 (A) and RTD-C7 (B). The HPLC traces of the corresponding linear peptide thioester precursor TFA crudes are shown on the left. The thioester linear peptide is marked in each case with an arrow. Other minor peptide products are marked with an asterisk. These compounds were identified as partially oxidized linear peptides by ES-MS and disappear once the cyclization/folding is complete (right HPLC traces).

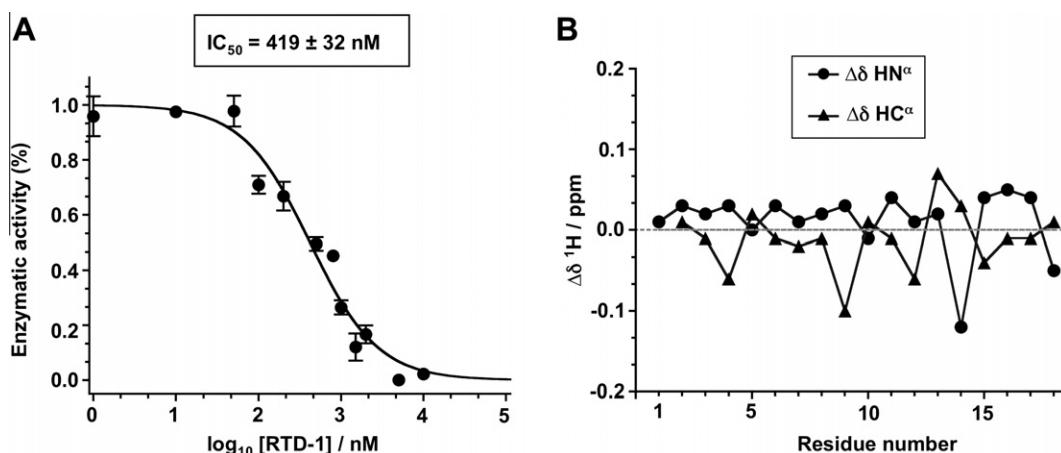


Figure 3. Characterization of synthetic RTD-1. (A) Inhibition assay of RTD-1 against anthrax lethal factor (LF). LF activity was measured in the presence of increasing concentrations of peptide RTD-1 and it was determined as the rate LF cleaves a FRET-based substrate.³² (B) Summary of the ¹H NMR assignments for the backbone protons: $\Delta\delta(^1\text{H})$ are the deviations in the chemical shifts of the main chain protons between the values obtained in this work for synthetic RTD-1 and those reported for RTD-1 at pH 6 (Table S1) and pH 4.5.²⁰

shifts of the assigned backbone amide and alpha protons (HN^α and HC^α) for synthetic RTD-1 were almost identical ($\leq 0.1 \text{ ppm}$) to those reported for RTD-1 at pH 6.0 (Fig. 3B and Table S1).³⁶ We also compared the chemical shifts of synthetic RTD-1 with those reported by Craik for RTD-1 in 10% MeCN at pH 4.5.²⁰ Again, no significant ($\leq 0.3 \text{ ppm}$) differences were found in the backbone amide protons. We also saw a uniform shift rather than variable changes for the backbone alpha protons ($\approx 0.2 \text{ ppm}$) at pH 6.0 (Fig. 3B and Table S1). Since backbone alpha protons usually reflect the secondary structure of the peptide backbone, this uniform offset in the resonances of the backbone HC^α protons can be attributed to the

different buffer conditions used in the two samples rather than changes in secondary structure. Altogether, these data confirm that synthetic RTD-1 adopts a native θ -defensin fold.

In summary, we have achieved the efficient synthesis of the θ -defensin RTD-1 using optimized Fmoc-based SPPS in combination with one-pot NCL-driven backbone cyclization with concomitant folding. We have shown that the GSH-induced cyclization/folding of linear precursor RTD-C7 is extremely clean not requiring the purification of the final cyclic folded θ -defensin. To our knowledge this represents the first chemical synthesis of RTD-1 using a chemical approach that is extremely efficient therefore requiring mini-

mal or no purification at all and therefore fully compatible with the high throughput production of chemical libraries using this scaffold. These libraries could be used for the optimization of the already known properties of θ -defensins or for the introduction of novel biological activities into this backbone cyclized polypeptide scaffold.

Acknowledgments

This work was supported by National Institutes of Health Research Grant R01-GM090323 (J.A.C.) and NIH Grant 5R01GM085006-02 (A.S.); and by the Department of Defense Congressionally Directed Medical Research Program Grant PC09305 (J.A.C.).

Supplementary data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.bmcl.2012.02.080](https://doi.org/10.1016/j.bmcl.2012.02.080).

References and notes

1. Tang, Y. Q.; Yuan, J.; Osapay, G.; Osapay, K.; Tran, D.; Miller, C. J.; Ouellette, A. J.; Selsted, M. E. *Science* **1999**, 286, 498.
2. Tanabe, H.; Yuan, J.; Zaragoza, M. M.; Dandekar, S.; Henschen-Edman, A.; Selsted, M. E.; Ouellette, A. J. *Infect. Immun.* **2004**, 72, 1470.
3. Selsted, M. E.; Ouellette, A. J. *Nat. Immunol.* **2005**, 6, 551.
4. Lehrer, R. I. *Nat. Rev.* **2004**, 2, 727.
5. Selsted, M. E. *Curr. Protein Pept. Sci.* **2004**, 5, 365.
6. Nguyen, T. X.; Cole, A. M.; Lehrer, R. I. *Peptides* **2003**, 24, 1647.
7. Garcia, A. E.; Osapay, G.; Tran, P. A.; Yuan, J.; Selsted, M. E. *Infect. Immun.* **2008**, 76, 5883.
8. Stegemann, C.; Tsvetkova, E. V.; Aleshina, G. M.; Lehrer, R. I.; Kokryakov, V. N.; Hoffmann, R. *Rapid Commun. Mass Spectrom.* **2010**, 24, 599.
9. Cole, A. M.; Hong, T.; Boo, L. M.; Nguyen, T.; Zhao, C.; Bristol, G.; Zack, J. A.; Waring, A. J.; Yang, O. O.; Lehrer, R. I. *Proc. Natl. Acad. Sci. U.S.A.* **1813**, 2002, 99.
10. Leonova, L.; Kokryakov, V. N.; Aleshina, G.; Hong, T.; Nguyen, T.; Zhao, C.; Waring, A. J.; Lehrer, R. I. *J. Leukoc. Biol.* **2001**, 70, 461.
11. Tran, D.; Tran, P. A.; Tang, Y. Q.; Yuan, J.; Cole, T.; Selsted, M. E. *J. Biol. Chem.* **2002**, 277, 3079.
12. Tongaonkar, P.; Tran, P.; Roberts, K.; Schaal, J.; Osapay, G.; Tran, D.; Ouellette, A. J.; Selsted, M. E. *J. Leukoc. Biol.* **2011**, 89, 283.
13. Tran, D.; Tran, P.; Roberts, K.; Osapay, G.; Schaal, J.; Ouellette, A.; Selsted, M. E. *Antimicrob. Agents Chemother.* **2008**, 52, 944.
14. Yasin, B.; Wang, W.; Pang, M.; Cheshenko, N.; Hong, T.; Waring, A. J.; Herold, B. C.; Wagar, E. A.; Lehrer, R. I. *J. Virol.* **2004**, 78, 5147.
15. Penberthy, W. T.; Chari, S.; Cole, A. L.; Cole, A. M. *Cell. Mol. Life Sci.* **2011**, 68, 2231.
16. Munk, C.; Wei, G.; Yang, O. O.; Waring, A. J.; Wang, W.; Hong, T.; Lehrer, R. I.; Landau, N. R.; Cole, A. M. *AIDS Res. Hum. Retroviruses* **2003**, 19, 875.
17. Owen, S. M.; Rudolph, D.; Wang, W.; Cole, A. M.; Sherman, M. A.; Waring, A. J.; Lehrer, R. I.; Lal, R. B. *J. Pept. Res.* **2004**, 63, 469.
18. Gallo, S. A.; Wang, W.; Rawat, S. S.; Jung, G.; Waring, A. J.; Cole, A. M.; Lu, H.; Yan, X.; Daly, N. L.; Craik, D. J.; Jiang, S.; Lehrer, R. I.; Blumenthal, R. *J. Biol. Chem.* **2006**, 281, 18787.
19. Wang, W.; Mulakala, C.; Ward, S. C.; Jung, G.; Luong, H.; Pham, D.; Waring, A. J.; Kaznessis, Y.; Lu, W.; Bradley, K. A.; Lehrer, R. I. *J. Biol. Chem.* **2006**, 281, 32755.
20. Trabi, M.; Schirra, H. J.; Craik, D. J. *Biochemistry* **2001**, 40, 4211.
21. Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. *Science* **1994**, 266, 776.
22. Dawson, P. E.; Kent, S. B. *Annu. Rev. Biochem.* **2000**, 69, 923.
23. Camarero, J. A.; Muir, T. W. *Chem. Commun.* **1997**, 1997, 1369.
24. Camarero, J. A.; Muir, T. W. *J. Am. Chem. Soc.* **1999**, 121, 5597.
25. Camarero, J. A.; Pavel, J.; Muir, T. W. *Angew. Chem., Int. Ed.* **1998**, 37, 347.
26. Kimura, R. H.; Tran, A. T.; Camarero, J. A. *Angew. Chem., Int. Ed.* **2006**, 45, 973.
27. Ingenito, R.; Bianchi, E.; Fattori, D.; Pessi, A. *J. Am. Chem. Soc.* **1999**, 121, 11369.
28. Shin, Y.; Winans, K. A.; Backes, B. J.; Kent, S. B. H.; Ellman, J. A.; Bertozi, C. R. *J. Am. Chem. Soc.* **1999**, 121, 11684.
29. Austin, J.; Wang, W.; Puttamadappa, S.; Shekhtman, A.; Camarero, J. A. *Chembiochem* **2009**, 10, 2663.
30. Austin, J.; Kimura, R. H.; Woo, Y. H.; Camarero, J. A. *Amino Acids* **2010**, 38, 1313.
31. Contreras, J.; Elnagar, A. Y. O.; Hamm-Alvarez, S.; Camarero, J. A. *J. Controlled Release* **2011**, 155, 134.
32. Kimura, R. H.; Steenblock, E. R.; Camarero, J. A. *Anal. Biochem.* **2007**, 369, 60.
33. Cheng, Y.; Prusoff, W. H. *Biochem. Pharmacol.* **1973**, 22, 3099.
34. Tonello, F.; Ascenzi, P.; Montecucco, C. *J. Biol. Chem.* **2003**, 278, 40075.
35. Kawakami, T.; Ohta, A.; Ohuchi, M.; Ashigai, H.; Murakami, H.; Suga, H. *Nat. Chem. Biol.* **2009**, 5, 888.
36. Gould, A.; Li, Y.; Majumder, S.; Garcia, A. E.; Carlsson, P.; Shekhtman, A.; Camarero, J. A. *Mol. Biosyst.* **2012**, 8, 1359.

Biological Synthesis of Circular Polypeptides*

Published, JBC Papers in Press, June 14, 2012, DOI 10.1074/jbc.R111.305508

Teshome L. Aboye[‡] and Julio A. Camarero^{‡§†}

From the Departments of [‡]Pharmacology and Pharmaceutical Sciences

and [§]Chemistry, University of Southern California, Los Angeles,

California 90033

Here, we review the use of different biochemical approaches for biological synthesis of circular or backbone-cyclized proteins and peptides. These methods allow the production of circular polypeptides either *in vitro* or *in vivo* using standard recombinant DNA expression techniques. Protein circularization can significantly impact protein engineering and research in protein folding. Basic polymer theory predicts that circularization should lead to a net thermodynamic stabilization of a folded protein by reducing the entropy associated with the unfolded state. Protein cyclization also provides a valuable tool for exploring the effects of topology on protein folding kinetics. Furthermore, the biological production of cyclic polypeptides makes possible the production of cyclic polypeptide libraries. The generation of such libraries, which was previously restricted to the domain of synthetic chemists, now offers biologists access to highly diverse and stable molecular libraries for probing protein structure and function.

Protein engineering is usually defined as the modification of the primary sequence of a protein to change in a defined way its three-dimensional structure and biological function. Site-specific mutagenesis using either recombinant DNA technology or chemical synthesis has been successfully applied to numerous proteins. Another useful approach to protein engineering involves changing backbone topology. Backbone cyclization or circularization (*i.e.* the covalent linkage of the N-terminal amino and C-terminal carboxylic groups) is a powerful tool for the study and manipulation of protein structure and function. The introduction of a covalent bond between the N and C termini of a protein provides a more stable topological constraint than the use of disulfide bonds. Hence, in proteins whose N and C termini are close in the native structure (a surprisingly common feature in folded proteins, particularly in single-domain proteins (1)), backbone cyclization should stabilize the protein fold by reducing the backbone entropy associated with the unfolded state of a protein. Backbone cyclization also has the added benefit of making proteins more resistant to proteolytic degradation, in particular against exoproteases (2, 3), which should increase both the *in vitro* stability of industrial enzymes and the *in vivo* stability of proteins with interesting pharmacological properties. In addition, protein cyclization provides a

valuable tool for exploring fundamental questions in protein folding, in particular how topology affects the kinetics of protein folding (4).

Backbone peptide cyclization has also been widely used in bioorganic and medicinal chemistry to improve the biochemical and biophysical properties of flexible and labile peptides in the development of peptide-based drug candidates (2, 3). The cyclization of flexible linear peptides reduces their conformational freedom and creates constrained structural frameworks that often confer high receptor binding affinity, specificity, and enhanced stability (2, 3).

Backbone-cyclized polypeptides (ranging from small peptides to small proteins) are also found throughout nature from bacteria to animals serving in a variety of functional roles (5). Many of these are synthesized by multienzyme complexes such as cyclosporine from fungi and daptomycin from bacteria. However, an increasing number of circular polypeptides are being discovered that are ribosomally produced as precursors and then post-translationally modified by cyclization (5). For example, θ -defensins are broad-spectrum antimicrobial 18-residue circular peptides expressed in blood leukocytes and bone marrow of Old World monkeys and represent the only circular peptides found in mammals to date. Backbone-cyclized polypeptides are also found in plants. Sunflower trypsin inhibitor-1 (SFTI-1),² for example, is a bicyclic 14-residue peptide found in sunflower seeds with potent trypsin inhibitory activity. Cyclotides comprise a novel family of well folded small globular backbone-cyclized microproteins (\approx 30 residues long) found in plants. Naturally occurring cyclotides have been characterized with insecticidal, uterotonic, anti-HIV, antimicrobial, antitumor, and antihelminthic activities (see Refs. 6 and 7 for recent reviews) and have been reported to cross cell membranes (8, 9). These properties make cyclotides highly attractive for the development of novel peptide-based therapeutics (10, 11).

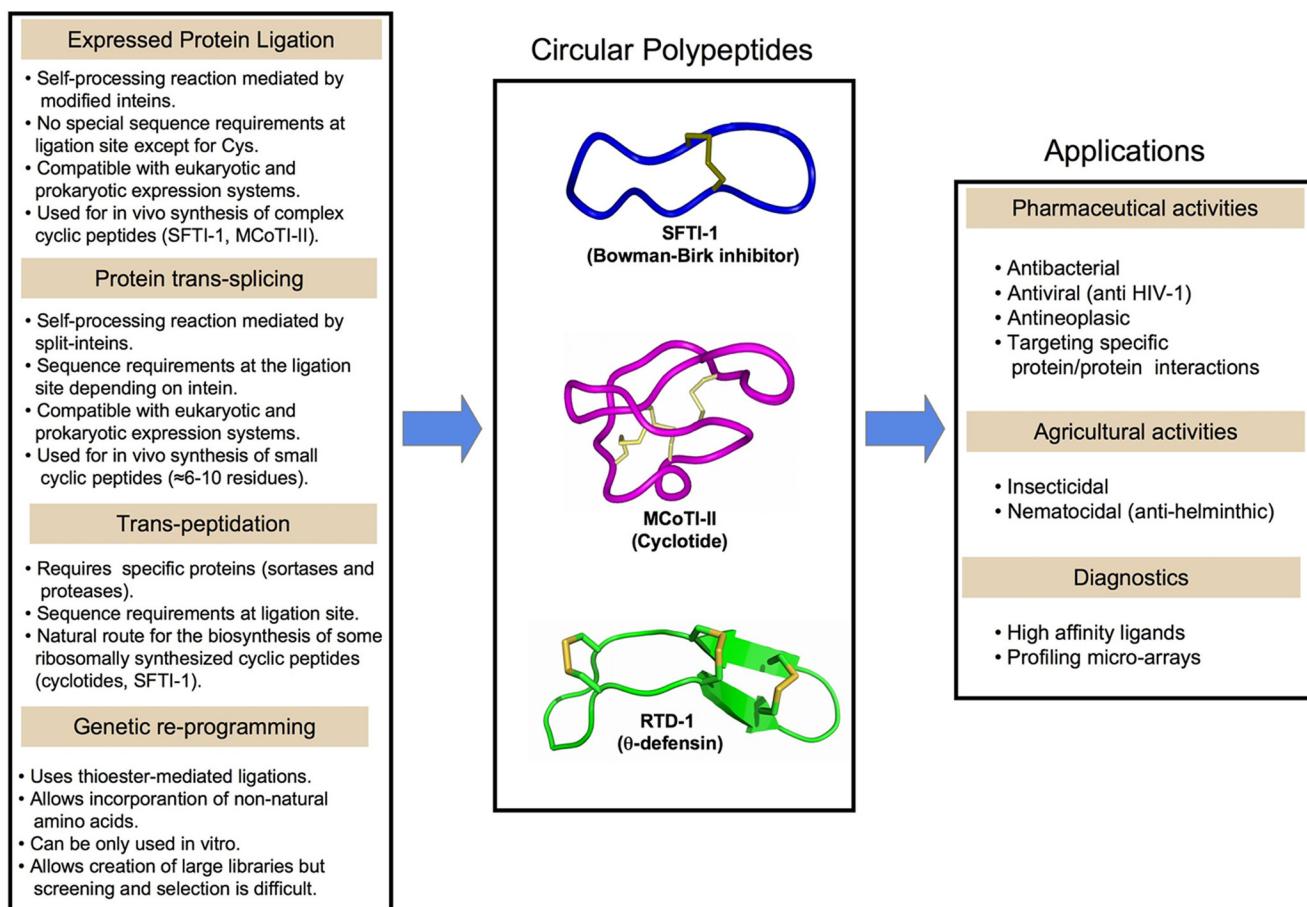
Despite the fact that the chemical synthesis of backbone-cyclized peptides and small protein domains has been well explored and a number of different approaches involving solid-phase or liquid-phase chemistry exist (12), recent advances in the fields of molecular biology and protein engineering have now made possible the biosynthesis of cyclic polypeptides using standard expression systems (13). The biological synthesis (*i.e.* the production of polypeptides mediated by ribosomal translation) of circular polypeptides offers many advantages over chemical synthesis. Although the chemical synthesis of polypeptides is limited to sequences no longer than 100–150 residues long, biological approaches do not present such limitations. Biological approaches also allow, using standard molecular biology tools, the generation of large combinatorial libraries of circular polypeptides that can be screened inside living cells for their ability to attenuate or inhibit cellular pro-

* This work was supported, in whole or in part, by National Institutes of Health Research Grant R01 GM090323 and Department of Defense Congressionally Directed Medical Research Program Grant PC09305 (to J. A. C.). This is the fifth article in the Thematic Minireview Series on Circular Proteins.

[†] To whom correspondence should be addressed. E-mail: jcamarero@usc.edu.

² The abbreviations used are: SFTI-1, sunflower trypsin inhibitor-1; EPL, expressed protein ligation; NCL, native chemical ligation; SH3, Src homology 3; PTS, protein trans-splicing; SML, sortase-mediated ligation; RTD-1, rhesus θ -defensin-1.

Biological methods



SCHEME 1. Summary of methods employed for biological production of backbone-cyclized peptides. Left panel, these approaches rely on the ribosomal synthesis of a polypeptide precursor that undergoes cyclization mediated by inteins or proteases. Middle panel, shown are three-dimensional structures of three examples of naturally occurring circular peptides with potential biotechnological value. MCoTi-II and SFTI-1 are powerful protease inhibitors and are attractive scaffolds for the introduction of other biological activities. RTD-1 is a θ -defensin that has strong antimicrobial and anti-inflammatory properties. All have been successfully expressed in biological systems (see Table 1). Right panel, the potential biotechnological applications of circular polypeptides, which include pharmacological, agrochemical, and diagnostic applications, are summarized. Molecular structures were generated using the structures of MCoTi-II (Protein Data Bank code 1IB9), SFTI-1 (code 1JBN), and RTD-1 (code 1HVZ).

cesses analogous to the way that the yeast two-hybrid technology works (14).

Different approaches have been described for the biological generation of circular polypeptides, including the use of expressed protein ligation (15), intein- and enzyme-mediated protein trans-splicing reactions (16), and genetic code reprogramming (17). In this minireview, we examine these methods as well as their potential applications, which are summarized in Scheme 1 and Table 1.

Backbone Cyclization Using Expressed Protein Ligation

Expressed protein ligation (EPL) (18) is an extension of native chemical ligation (NCL), a chemoselective ligation approach for the chemical synthesis of proteins (19). In this reaction, a fully unprotected peptide with an α -thioester at the C terminus reacts chemoselectively under neutral aqueous conditions with another unprotected peptide containing an N-terminal Cys residue to form a native peptide bond (19). It is also well established that when these two reactive groups are located in the same linear polypeptide precursor, the chemical ligation can proceed in an intramolecular fashion, therefore giving rise to a

backbone-cyclized polypeptide (20). This reaction has been successfully used for the chemical synthesis of circular peptides and small protein domains (20).

Recent advances in protein engineering, as well as the discovery of protein splicing, have also made possible the production of recombinant proteins containing an N-terminal Cys residue and/or the C-terminal α -thioester functionality. These important developments have made it possible to perform NCL reactions between recombinant and/or synthetic polypeptides. This technology, called EPL, has provided access to a multitude of chemically engineered proteins, including backbone-cyclized polypeptides (see Ref. 20 for a recent review).

The use of EPL-mediated backbone cyclization for the biosynthesis of circular polypeptides (Fig. 1A) was first reported by Camarero and Muir in 1999 (21) using the N-terminal SH3 domain of the c-Crk protein as a model system. In this work, the SH3 domain was fused to a modified *Saccharomyces cerevisiae* vacuolar membrane ATPase intein at the C terminus and to the MIEGRC motif (which contains a factor Xa proteolysis site) at the N terminus. The intein fusion construct was then expressed in *Escherichia coli*, purified, and cleaved with factor Xa, produc-

TABLE 1

Examples of backbone-cyclized polypeptides produced by ribosomal synthesis

These were produced using EPL, PTS, SML, genetic reprogramming, and protease-mediated ligation. The sizes of the polypeptides shown range from small peptides (eight residues) to proteins (263 residues).

Protein/peptide	No. of residues	Method	Application	Ref./source
c-Crk SH3	57	EPL	Stability Folding kinetics	21 49
TEM-1 β -lactamase	263	EPL PTS	Stability	24 34
Kalata B1	29	EPL	Natural product	27
MCoTI-I/II	34	EPL	Natural product Library scaffold	28 29
Cryptdin-4	32	Trypsin-mediated cyclization	Natural product	42
SFTI-1	14	EPL	Stability and improved activity	31
		EPL	Natural product and library scaffold	30
		Trypsin-mediated cyclization	Natural product and mutants	9,41
RTD-1	18	Genetic reprogramming	Natural product and library scaffold	46
Pseudostellarin F	8	PTS	Natural product	16
HIV protease inhibitors	6	PTS	Library scaffold and screening	33
GFP	238	EPL	Stability	25
IFN α 2	188	SML	Thermal and serum stability	38
Histatin-1	38	SML	Improve biological activity	39
				40

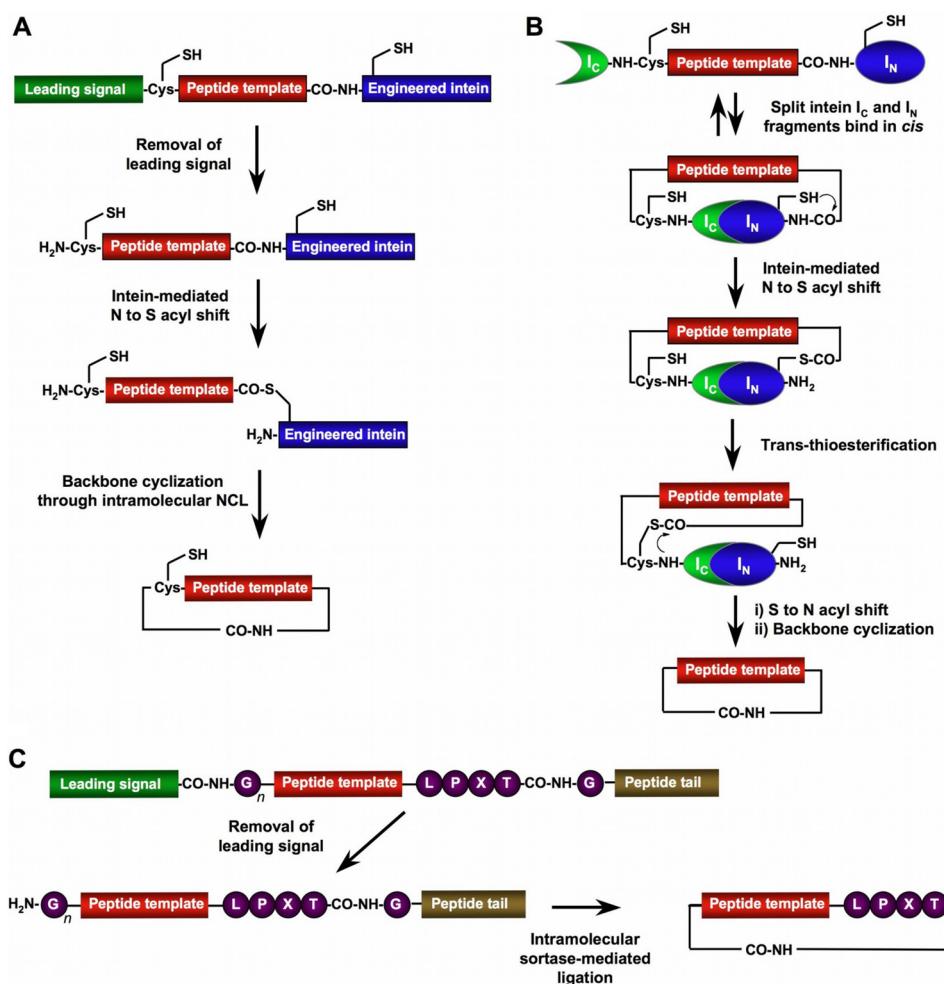


FIGURE 1. Three most common approaches for biological production of backbone-cyclized polypeptides using EPL (A), PTS (B), and SML (C). I_C, C-intein; I_N, N-intein.

ing an N-terminal Cys-containing SH3-intein fusion protein, which spontaneously reacted in an intramolecular fashion to yield the fully active circular SH3 domain (21). The cyclization reaction was extremely efficient and clean. It is also worth men-

tioning that in contrast to standard EPL protocols, which usually require denaturing conditions and the presence of thiol cofactors (18), the cyclization reaction was performed under physiological conditions at neutral pH in the absence of thiol

cofactors. This was attributed to the close proximity of the α -thioester and N-terminal Cys residue in the folded state of the SH3 domain, which increased the local concentration of the reacting groups, thus facilitating the cyclization reaction. This effect has also been found in the cyclization of other small protein domains (22, 23). A similar approach has also been used by Iwai and Plückthun for the biosynthesis of circular β -lactamase (24) and GFP (25). The resulting circular proteins were biologically active and found to be more resistant against thermal denaturation (24, 25).

Inspired by the high efficiency of the *in vitro* cyclization of the SH3 domain (21), Camarero *et al.* (26) used a similar approach to explore the cyclization of the SH3 domain inside living cells. This work provided the first demonstration of polypeptide chemical ligation reactions performed inside the complex cytoplasmic environment of a living cell. In-cell production of circular polypeptides makes possible the generation of genetically encoded libraries that can be assayed in the intracellular milieu for rapid screening of biologically active compounds.

More recently, our laboratory has used this approach for the recombinant production of cyclotides using standard bacterial expression systems (27, 28). Cyclotides are small globular microproteins with a head-to-tail cyclized backbone, which is further stabilized by three disulfide bonds forming a cystine knot. This cystine knot motif gives cyclotides exceptional resistance to physical, chemical, and biological degradation. Our group has also pioneered the use of EPL-mediated backbone cyclization for the biosynthesis of fully folded cyclotides inside bacterial cells using standard expression systems (28, 29). We have recently used this approach for the generation of a small library based on the cyclotide MCoTI-I, in which all of the residues except Cys and those located in loop 6 were replaced by different residues (29). The resulting library was screened for biological activity and the ability to adopt a cyclotide fold, indicating that most of the residues in the cyclotide scaffold could be replaced without significantly affecting the ability of the resulting sequences to fold correctly (29). These results highlight the robustness of the cyclotide scaffold for generating large genetically encoded libraries of cyclotides that can be screened in-cell using high throughput techniques.

EPL-mediated backbone cyclization of polypeptides has also been recently used for the biosynthesis of the Bowman-Birk inhibitor SFTI-1 (30). The biosynthesis of other cyclic peptides such as backbone-cyclized α -defensins (31) and naturally occurring θ -defensins (50) has also been produced in our laboratory using bacterial expression systems, thus showing the general applicability of this method for the production of naturally occurring or engineered circular polypeptides.

Backbone Cyclization Using Intein-mediated Protein Trans-splicing

Another method to generate cyclic peptides *in vivo* is by protein trans-splicing (PTS) (Fig. 1B). This approach utilizes a self-processing intein that is split into two fragments, an N-intein and a C-intein. This technique was first reported by Benkovic and co-workers (32) using the naturally occurring *Synechocystis*

sp. PCC6803 DnaE split intein. PTS is a post-translational modification similar to protein splicing with the difference that the intein self-processing domain is split into two fragments (N-intein and C-intein). The split intein fragments are not active individually; however, they can bind to each other with high specificity under appropriate conditions to form an active protein-splicing or intein domain in *trans*. By rearranging the order of the intein fragments through fusing the N- and C-intein fragments to the C and N termini of the polypeptide to be cyclized, the trans-splicing reaction results in the formation of a backbone-cyclized polypeptide (Fig. 1B). This methodology, often called SICLOPPS (split intein circular ligation of proteins and peptides), has been successfully used for the generation of several naturally occurring cyclic peptides as well as large genetically encoded libraries of small cyclic peptides (16). More recently, Schultz and co-workers (33) reported the use of this technology in combination with nonsense codon suppressor tRNA technology to build large libraries of cyclic hexapeptides using an expanded set of non-natural amino acid building blocks. These libraries were then used to screen for an inhibitor of HIV protease using a selection based on cellular viability (33). This was accomplished by linking cellular viability to the activity of HIV protease. In this particular case, the HIV protease cleavage site was introduced into the flexible hinge region of the tetracycline resistance gene (Tn10-HIV). The introduction of this sequence does not affect significantly the biological activity of Tn10-HIV; however, when it is coexpressed with HIV protease, resistance to tetracycline is abolished. Hence, cells coexpressing Tn10-HIV and HIV protease would be able to grow in medium containing tetracycline only if the HIV protease was inhibited. The authors used this elegant approach for screening surviving cells coexpressing HIV/Tn10-HIV and transformed with a genetically encoded library of cyclic peptides. Under these conditions, only cells expressing cyclic peptides able to inhibit HIV protease were able to grow to saturation (33). The authors were able to select different cyclic hexapeptides containing the non-natural amino acid *p*-benzoylphenylalanine, which contains a keto group and is able to inhibit HIV with IC_{50} values in the low micromolar range (33). These interesting findings suggest that the use of an expanded genetic code using novel chemical functionalities not present in nature such as a keto group could confer an evolutionary advantage in response to selective pressure.

PTS has also been used for the generation of circular proteins (32, 34). In a recent study, an artificially split intein DnaB mini-intein from the *Synechocystis* sp. PCC6803 strain was used for the cyclization of the TEM-1 β -lactamase in the bacterial periplasm; this was accomplished by adding TEM-1 β -lactamase periplasm export signal peptide to the split intein precursor (34). The same authors also used this approach for the production of large libraries of small circular peptides. It was estimated that ~50% of the combinatorial peptides were able to cyclize efficiently.

Although the use of PTS is a good alternative for the biosynthesis of backbone-cyclized polypeptides and biologically encoded libraries, it should be noted, however, that it requires the presence of specific amino acid residues at both intein-extein junctions for efficient protein splicing to occur (16). To

resolve this, new or engineered split inteins that are more tolerant with non-native extein-intein junction sequences are required. In contrast, the only sequence requirement for the use of EPL or intein-mediated backbone cyclization is the presence of an N-terminal Cys. Model studies have also shown that all of the 20 natural amino acids located at the C terminus of a polypeptide thioester can support ligation (35). It is worth noting, however, that the speed rate of thioester-mediated ligations depends on the nature of the amino acid at the C terminus, *e.g.* Gly and Ala residues react faster, whereas β -branched residues react slower and produce product in lower yields. Furthermore, the engineered inteins most commonly used for the generation of polypeptide α -thioesters have also been shown to be compatible with most amino acids upstream of the cleavage site (20).

Backbone Cyclization Using Protease-catalyzed Transpeptidation

It is well known that prokaryotes use protease-catalyzed protein splicing to attach proteins to peptidoglycan in their cell wall envelope. Sortases are transpeptidase enzymes found in most Gram-positive bacteria that are specialized in this task. Among the different bacterial sortases isolated so far, the *Staphylococcus aureus* sortase A (SrtA) has been by far the most widely used for protein engineering (36). SrtA recognizes peptide substrates containing the sequence LPXTG (where X can be any residue), cleaving the peptide bond between the Thr and Gly residues. The cleavage is catalyzed by an active site Cys residue that generates a transient thioacyl-enzyme intermediate. This thioester-reactive species then reacts with the α -amino group of oligoglycine (pentaglycine in the case of SrtA) substrates to produce ligated products. In the absence of oligoglycine substrates, the thioacyl-enzyme intermediate is hydrolyzed.

Backbone-cyclized polypeptides can be obtained by using an intramolecular version of the sortase-mediated ligation (SML) (Fig. 1C). The first example of polypeptide cyclization by SML was reported by Boder and co-workers (37). In this work, a GFP construct containing the SrtA recognition sequences at the N and C termini, respectively, was backbone-cyclized through SML. Ploegh and co-workers (38) have also recently used a similar approach to obtain a circular version of GFP using SrtA. SrtA has also been used recently for the backbone cyclization of several cytokines to improve their stability and serum half-life (39). In this elegant work, SML was used in tandem for the introduction of an aminoxyacetic acid moiety at the termini of several cytokines, which were then backbone-cyclized and site-specifically PEGylated through the aminoxyacetic acid moiety for increased stability in serum (39). SrtA has also been used recently to produce a backbone-cyclized version of histatin-1, a 38-mer salivary polypeptide with antimicrobial and wound-healing activities (40). According to the authors, the SrtA-mediated cyclization of histatin-1 provided better yields than chemical synthesis. It should be highlighted, however, that in this case, the synthesis was performed entirely using a solid-phase approach, *i.e.* the synthesis of the linear precursor, as well as its cyclization, was carried out entirely on a solid support using standard chemical coupling reagents. This type of

approach has been shown to lead to lower yields for the cyclization of long polypeptides and small protein domains (22, 23). In contrast, the use of chemoselective ligations such as NCL (which uses the same type of chemoselective ligation as EPL) (Fig. 1A) has proven to give excellent results for the backbone cyclization of polypeptides and small protein domains (22, 23).

Although the use of SML for the production of cyclic polypeptides shows great promise, it has not been shown to work in the cell yet, which limits its use for the generation of cellular libraries of circular polypeptides for genetic experiments. Moreover, the actual approach requires the use of rather large sortase recognition sequences (LPXTG₅ in the case of SrtA) (Fig. 1C) that remain incorporated in the final cyclic polypeptide, which could affect the applicability of this approach for some applications.

More promising is the use of “traceless” protease-catalyzed protein splicing for cyclization of the polypeptide backbone. For example, Craik and co-workers (41) have shown that trypsin can be used for the efficient backbone cyclization of the cyclic peptide SFTI-1. A similar approach has also been used for the cyclization of the cyclotides MCoTI-I and MCoTI-II *in vitro* (42). MCoTI cyclotides and SFTI-1 are potent trypsin inhibitors, with K_i values in the picomolar and nanomolar range, respectively. In both cases, the proteolytic enzyme trypsin was able to resynthesize the peptide bond between the P1 and P1' residues to form the native circular polypeptides (41, 42). Although this approach has not yet been tested *in vivo*, there is mounting evidence that protease-catalyzed protein splicing is used in the biosynthesis of these naturally occurring circular polypeptides in plants. Recent studies suggest that asparaginyl endopeptidases are involved in the biosynthesis of cyclotides and SFTI-1 through an asparaginyl endopeptidase-mediated transpeptidation step (43, 44).

Backbone Cyclization Using Genetic Code Reprogramming

Genetic code reprogramming is an emerging new method that allows the production *in vitro* of ribosomally expressed polypeptides containing non-proteinogenic amino acids (45). The key to this approach is the use of custom-made reconstituted *in vitro* translation systems and, more importantly, the ability to employ enzymes that are able to load specific amino acids into designed tRNAs. Among the different methods to accomplish the task is the use of an RNA-based enzyme, flexizyme, which shows great promise (45). Flexizyme is a ribozyme that catalyzes the transfer of practically any amino acid activated as a *p*-nitrophenyl ester into any given tRNA.

Suga and co-workers (46) have recently shown that this method can also be used for the generation of ribosomally translated backbone-cyclized polypeptides in cell-free expression systems. This was accomplished by the incorporation of the cysteinyl-prolinyl glycolic acid ester moiety at the C terminus of ribosomally produced polypeptides by reprogramming glycolic acid into a codon using the flexizyme system. Cysteinyl-prolinyl esters can rapidly rearrange at physiological pH through the formation of a diketopiperazine intermediate to produce a C-terminal thioester (46). This post-translationally formed thioester can then react in an intramolecular fashion

with an N-terminal Cys residue through NCL in the same way that EPL works (Fig. 1A). The incorporation of an N-terminal Cys into polypeptides generated *in vitro* can be easily accomplished by adding peptide deformylase and methionine aminopeptidase to the *in vitro* translation system. This approach has been used for the generation of numerous backbone-cyclized peptides, including SFTI-1 and rhesus θ -defensin-1 (RTD-1) (46).

Suga and co-workers (46) have also shown that this strategy can be used for the generation of peptide libraries. However, due to the circular nature of the peptides produced, it is difficult to genetically encode (e.g. using an RNA-display approach) the different members of a particular library generated by this technique. This makes the screening of such libraries a very challenging task, which limits the complexity of the libraries that can be analyzed using this method. Another limitation of this approach is that production yields are usually low, therefore limiting the utility of this method for screening purposes.

Summary and Outlook

In this minireview, we have briefly examined the different biochemical approaches available for the biological production of backbone-cyclized polypeptides ranging from small peptides to proteins. These include the use of protein splicing catalyzed by inteins (16) or proteases (41, 42) as well as thioester-based ligation using either modified inteins (27, 29, 30) or genetic code reprogramming (46). The ability to produce circular polypeptides using the tools of molecular biology has multiple biotechnological applications (Scheme 1). For example, it makes possible the generation of large libraries of cyclic polypeptides that can be screened for biological activity. Several of the methods discussed also allow the production of libraries inside living cells (29, 30, 33). This makes possible the rapid screening and selection of novel sequences that can modulate or inhibit the biological activities of particular biomolecular targets. This approach allows the use of molecular evolution and cell-based assays for the rapid selection of novel sequences with optimal biological activity. For example, Schultz and co-workers (33) have recently used PTS in combination with nonsense codon suppressor tRNA technology to build large libraries of cyclic hexapeptides incorporating non-natural amino acids inside bacterial cells that were later screened for activity against the HIV protease. Although the biological activities for the selected circular peptides were rather modest (IC_{50} values in the low micromolar range), this could be attributed to the small size of the cyclic peptide scaffold used in this work (only six residues). This could be easily addressed by using a cyclic polypeptide template incorporating more than one loop for potential interactions in the same way that antibody complementarity-determining regions work. Within this context, our group has pioneered the use of EPL-mediated backbone cyclization for the biosynthesis of cyclotide-based libraries inside living cells for the high throughput selection of novel biological activities. Our group is using the cyclotide MCoTI-I as a molecular scaffold, which has up to five hypervariable loops available for the design of complex libraries that can be used for molecular evolution experiments using cell-based screening. MCoTI cyclotides are nontoxic and are powerful trypsin inhibitors, with IC_{50} values

in the low picomolar range, which demonstrates the potential of this scaffold for selecting novel sequences that target protein-protein interactions with very high affinities. Moreover, the cyclic cystine knot motif found in cyclotides gives them enhanced thermal, chemical, and biological stabilities. All of these features make the cyclotide scaffold an ideal molecular framework for the development of novel peptide-based therapeutics.

In this context, it is worth noting that new critical information on the origin, evolution, and processing of cyclotides as well as other circular polypeptides found in plants is beginning to provide insight on how these polypeptides are post-translationally cyclized in plants. A complete elucidation of the mechanisms associated with the ribosomal biosynthesis of circular polypeptides in plants should make possible the production of pharmacologically useful circular polypeptides in genetically engineered plants in the near future.

The production of circular polypeptides using standard biological expression systems has also made possible the introduction of NMR active isotopes (^{15}N and/or ^{13}C) in a very inexpensive fashion (47). This should facilitate the use of NMR to study structure-activity relationships of any potential circular polypeptide and its molecular target (47). The recent development of in-cell NMR using sequential labeling approaches (48) could be easily interfaced for the in-cell screening of genetically encoded libraries of circular polypeptides.

In summary, the development of novel biochemical approaches for the biological synthesis of circular polypeptides is opening new and exciting opportunities that previously were available only to synthetic chemists. The production of complex libraries inside living cells using standard recombinant techniques offers biologists a unique opportunity to perform experiments in molecular evolution that will unravel genomic information encoding complex biochemical pathways and protein interaction networks. The use of highly stable circular polypeptide scaffolds provides an ideal molecular tool for this challenging task.

REFERENCES

- Thornton, J. M., and Sibanda, B. L. (1983) Amino- and carboxyl-terminal regions in globular proteins. *J. Mol. Biol.* **167**, 443–460
- Clark, R. J., Fischer, H., Dempster, L., Daly, N. L., Rosengren, K. J., Nevin, S. T., Meunier, F. A., Adams, D. J., and Craik, D. J. (2005) Engineering stable peptide toxins by means of backbone cyclization: stabilization of the α -conotoxin MII. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 13767–13772
- Clark, R. J., Akcan, M., Kaas, Q., Daly, N. L., and Craik, D. J. (2012) Cyclization of conotoxins to improve their biopharmaceutical properties. *Toxicon* **59**, 446–455
- Plaxco, K. W., Simons, K. T., Ruczinski, I., and Baker, D. (2000) Topology, stability, sequence, and length: defining the determinants of two-state protein folding kinetics. *Biochemistry* **39**, 11177–11183
- Craik, D. J. (2006) Chemistry. Seamless proteins tie up their loose ends. *Science* **311**, 1563–1564
- Daly, N. L., Rosengren, K. J., and Craik, D. J. (2009) Discovery, structure, and biological activities of cyclotides. *Adv. Drug Deliv. Rev.* **61**, 918–930
- Garcia, A. E., and Camarero, J. A. (2010) Biological activities of natural and engineered cyclotides, a novel molecular scaffold for peptide-based therapeutics. *Curr. Mol. Pharmacol.* **3**, 153–163
- Contreras, J., Elnagar, A. Y., Hamm-Alvarez, S. F., and Camarero, J. A. (2011) Cellular uptake of cyclotide MCoTI-I follows multiple endocytic pathways. *J. Controlled Release* **155**, 134–143

9. Cascales, L., Henriques, S. T., Kerr, M. C., Huang, Y. H., Sweet, M. J., Daly, N. L., and Craik, D. J. (2011) Identification and characterization of a new family of cell-penetrating peptides. *Cyclic cell-penetrating peptides*. *J. Biol. Chem.* **286**, 36932–36943
10. Henriques, S. T., and Craik, D. J. (2010) Cyclotides as templates in drug design. *Drug Discov. Today* **15**, 57–64
11. Jagadish, K., and Camarero, J. A. (2010) Cyclotides, a promising molecular scaffold for peptide-based therapeutics. *Biopolymers* **94**, 611–616
12. White, C. J., and Yudin, A. K. (2011) Contemporary strategies for peptide macrocyclization. *Nat. Chem.* **3**, 509–524
13. Sancheti, H., and Camarero, J. A. (2009) “Splicing up” drug discovery. Cell-based expression and screening of genetically encoded libraries of backbone-cyclized polypeptides. *Adv. Drug Deliv. Rev.* **61**, 908–917
14. Suter, B., Kittanakom, S., and Stagljar, I. (2008) Two-hybrid technologies in proteomics research. *Curr. Opin. Biotechnol.* **19**, 316–323
15. Kimura, R., and Camarero, J. A. (2005) Expressed protein ligation: a new tool for the biosynthesis of cyclic polypeptides. *Protein Pept. Lett.* **12**, 789–794
16. Tavassoli, A., and Benkovic, S. J. (2007) Split intein-mediated circular ligation used in the synthesis of cyclic peptide libraries in *E. coli*. *Nat. Protoc.* **2**, 1126–1133
17. Ohta, A., Yamagishi, Y., and Suga, H. (2008) Synthesis of biopolymers using genetic code reprogramming. *Curr. Opin. Chem. Biol.* **12**, 159–167
18. Muir, T. W. (2003) Semisynthesis of proteins by expressed protein ligation. *Annu. Rev. Biochem.* **72**, 249–289
19. Dawson, P. E., and Kent, S. B. (2000) Synthesis of native proteins by chemical ligation. *Annu. Rev. Biochem.* **69**, 923–960
20. Berrade, L., and Camarero, J. A. (2009) Expressed protein ligation: a resourceful tool to study protein structure and function. *Cell. Mol. Life Sci.* **66**, 3909–3922
21. Camarero, J. A., and Muir, T. W. (1999) Biosynthesis of a head-to-tail cyclized protein with improved biological activity. *J. Am. Chem. Soc.* **121**, 5597–5598
22. Camarero, J. A., Pavel, J., and Muir, T. W. (1998) *Angew. Chem. Int. Ed. Engl.* **37**, 347–349
23. Camarero, J. A., Cotton, G. J., Adeva, A., and Muir, T. W. (1998) Chemical ligation of unprotected peptides directly from a solid support. *J. Pept. Res.* **51**, 303–316
24. Iwai, H., and Plückthun, A. (1999) Circular β -lactamase: stability enhancement by cyclizing the backbone. *FEBS Lett.* **459**, 166–172
25. Iwai, H., Lingel, A., and Plückthun, A. (2001) Cyclic green fluorescent protein produced *in vivo* using an artificially split PI-Pful intein from *Pyrococcus furiosus*. *J. Biol. Chem.* **276**, 16548–16554
26. Camarero, J. A., Fushman, D., Cowburn, D., and Muir, T. W. (2001) Peptide chemical ligation inside living cells: *in vivo* generation of a circular protein domain. *Bioorg. Med. Chem.* **9**, 2479–2484
27. Kimura, R. H., Tran, A. T., and Camarero, J. A. (2006) Biosynthesis of the cyclotide kalata B1 by using protein splicing. *Angew. Chem. Int. Ed. Engl.* **45**, 973–976
28. Camarero, J. A., Kimura, R. H., Woo, Y. H., Shekhtman, A., and Cantor, J. (2007) Biosynthesis of a fully functional cyclotide inside living bacterial cells. *ChemBioChem* **8**, 1363–1366
29. Austin, J., Wang, W., Puttamadappa, S., Shekhtman, A., and Camarero, J. A. (2009) Biosynthesis and biological screening of a genetically encoded library based on the cyclotide MCoTI-I. *ChemBioChem* **10**, 2663–2670
30. Austin, J., Kimura, R. H., Woo, Y. H., and Camarero, J. A. (2010) *In vivo* biosynthesis of an Ala scan library based on the cyclic peptide SFTI-1. *Amino Acids* **38**, 1313–1322
31. Garcia, A. E., Tai, K. P., Puttamadappa, S. S., Shekhtman, A., Ouellette, A. J., and Camarero, J. A. (2011) Biosynthesis and antimicrobial evaluation of backbone-cyclized α -defensins. *Biochemistry* **50**, 10508–10519
32. Scott, C. P., Abel-Santos, E., Wall, M., Wahnon, D. C., and Benkovic, S. J. (1999) Production of cyclic peptides and proteins *in vivo*. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 13638–13643
33. Young, T. S., Young, D. D., Ahmad, I., Louis, J. M., Benkovic, S. J., and Schultz, P. G. (2011) Evolution of cyclic peptide protease inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 11052–11056
34. Deschuyteneer, G., Garcia, S., Michiels, B., Baudoux, B., Degand, H., Morisomme, P., and Soumilion, P. (2010) Intein-mediated cyclization of randomized peptides in the periplasm of *Escherichia coli* and their extracellular secretion. *ACS Chem. Biol.* **5**, 691–700
35. Hackeng, T. M., Griffin, J. H., and Dawson, P. E. (1999) Protein synthesis by native chemical ligation: expanded scope by using straightforward methodology. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 10068–10073
36. Tsukiji, S., and Nagamune, T. (2009) Sortase-mediated ligation: a gift from Gram-positive bacteria to protein engineering. *ChemBioChem* **10**, 787–798
37. Parthasarathy, R., Subramanian, S., and Boder, E. T. (2007) Sortase A as a novel molecular “stapler” for sequence-specific protein conjugation. *Bioconjug. Chem.* **18**, 469–476
38. Antos, J. M., Popp, M. W., Ernst, R., Chew, G. L., Spooner, E., and Ploegh, H. L. (2009) A straight path to circular proteins. *J. Biol. Chem.* **284**, 16028–16036
39. Popp, M. W., Dougan, S. K., Chuang, T. Y., Spooner, E., and Ploegh, H. L. (2011) Sortase-catalyzed transformations that improve the properties of cytokines. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 3169–3174
40. Bolscher, J. G., Oudhoff, M. J., Nazmi, K., Antos, J. M., Guimaraes, C. P., Spooner, E., Haney, E. F., Garcia Vallejo, J. J., Vogel, H. J., van't Hof, W., Ploegh, H. L., and Veerman, E. C. (2011) Sortase A as a tool for high yield histatin cyclization. *FASEB J.* **25**, 2650–2658
41. Marx, U. C., Korsinczky, M. L., Schirra, H. J., Jones, A., Condie, B., Otvos, L., Jr., and Craik, D. J. (2003) Enzymatic cyclization of a potent Bowman-Birk protease inhibitor, sunflower trypsin inhibitor-1, and solution structure of an acyclic precursor peptide. *J. Biol. Chem.* **278**, 21782–21789
42. Thongyoo, P., Roqué-Rosell, N., Leatherbarrow, R. J., and Tate, E. W. (2008) Chemical and biomimetic total syntheses of natural and engineered MCoTI cyclotides. *Org. Biomol. Chem.* **6**, 1462–1470
43. Gillon, A. D., Saska, I., Jennings, C. V., Guarino, R. F., Craik, D. J., and Anderson, M. A. (2008) Biosynthesis of circular proteins in plants. *Plant J.* **53**, 505–515
44. Mylne, J. S., Colgrave, M. L., Daly, N. L., Chanson, A. H., Elliott, A. G., McCallum, E. J., Jones, A., and Craik, D. J. (2011) Albumins and their processing machineries are hijacked for cyclic peptides in sunflower. *Nat. Chem. Biol.* **7**, 257–259
45. Goto, Y., Katoh, T., and Suga, H. (2011) Flexizymes for genetic code reprogramming. *Nat. Protoc.* **6**, 779–790
46. Kawakami, T., Ohta, A., Ohuchi, M., Ashigai, H., Murakami, H., and Suga, H. (2009) Diverse backbone-cyclized peptides via codon reprogramming. *Nat. Chem. Biol.* **5**, 888–890
47. Puttamadappa, S. S., Jagadish, K., Shekhtman, A., and Camarero, J. A. (2010) Backbone dynamics of cyclotide MCoTI-I free and complexed with trypsin. *Angew. Chem. Int. Ed. Engl.* **49**, 7030–7034
48. Burz, D. S., Dutta, K., Cowburn, D., and Shekhtman, A. (2006) Mapping structural interactions using in-cell NMR spectroscopy (STINT-NMR). *Nat. Methods* **3**, 91–93
49. Camarero, J. A., Fushman, D., Sato, S., Giriat, I., Cowburn, D., Raleigh, D. P., and Muir, T. W. (2001) Rescuing a destabilized protein fold through backbone cyclization. *J. Mol. Biol.* **308**, 1045–1062
50. Gould, A., Li, Y., Majumder, S., Garcia, A. E., Carlsson, P., Shekhtman, A., and Camarero, J. A. (2012) Recombinant production of rhesus θ -defensin-1 using a bacterial expression system. *Mol. Biosyst.* **8**, 1359–1365



Expression of Fluorescent Cyclotides using Protein Trans-Splicing for Easy Monitoring of Cyclotide–Protein Interactions**

Krishnappa Jagadish, Radhika Borra, Vanessa Lacey, Subhabrata Majumder, Alexander Shekhtman, Lei Wang, and Julio A. Camarero*

Cyclotides are fascinating natural plant micro-proteins ranging from 28 to 37 amino acid residues long and exhibit various biological functions including anti-microbial, insecticidal, cytotoxic, antiviral (against HIV), protease inhibitory, and hormone-like activities.^[1–4] They share a unique head-to-tail circular knotted topology of three disulfide bridges; one disulfide penetrates through a macrocycle formed by the other two disulfides, thereby interconnecting the peptide backbone to form what is called a cystine knot topology (Figure 1). This cyclic cystine knot (CCK) framework gives the cyclotides exceptional rigidity,^[5] resistance to thermal and chemical denaturation, and enzymatic stability against degradation.^[4,6] In fact, some cyclotides have been shown to be orally bioavailable. For example, the first cyclotide to be discovered, kalata B1, was found to be an orally effective uterotonic,^[7] and other cyclotides have been shown to cross the cell membrane through macropinocytosis.^[8–10] All of these features make cyclotides ideal substrates for drug development.^[11–14]

Cyclotides have been isolated from plants in the *Rubiaceae*, *Violaceae*, *Cucurbitaceae*,^[4,15] and, most recently, *Fabaceae* families.^[16–18] Around 200 different cyclotide sequences have been reported in the literature,^[19,20] although it has been estimated that approximately 50 000 cyclotides may exist.^[21,22] Despite sequence diversity, all cyclotides share the same CCK motif (Figure 1A). Hence, these micro-proteins can be

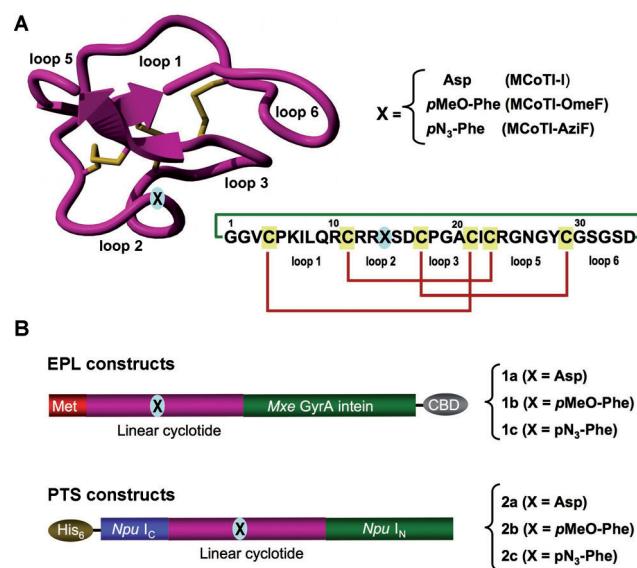


Figure 1. A) Tertiary structure of the cyclotide MCoTi-II (PDB code: 1IB9)^[52] and the primary sequence of cyclotides used in this study. The backbone-cyclized peptide (connecting bond shown in green) is stabilized by the three disulfide bonds (shown in red). B) Intein precursors used for the expression of cyclotides. A non-natural amino acid (X) was introduced at position 14, which is in the middle of loop 2.

considered natural combinatorial peptide libraries that are structurally constrained by the cystine-knot scaffold^[2] and head-to-tail cyclization, and in which, with the exception of the strictly conserved cysteine residues comprising the cysteine-knot, hypermutation of essentially all residues is permitted. Cyclotides can be chemically synthesized, thereby permitting the introduction of specific chemical modifications or biophysical probes.^[14,23–25] More importantly, cyclotides can now be biosynthesized in bacterial cells using a biomimetic approach that involves the use of modified protein splicing units.^[26–28] These characteristics make them ideal substrates for the production of genetically-encoded libraries based on the cyclotide framework. These cell-based libraries allow for in-cell molecular-evolution strategies to enable the generation and high-throughput selection of compounds with optimal binding and inhibitory characteristics. In contrast to chemically generated libraries, genetically encoded libraries enable the facile generation and screening of very large combinatorial libraries of molecules.

The genetic code of most organisms encodes only 20 canonical amino acid building blocks, with the rare exceptions

[*] Dr. K. Jagadish, Dr. R. Borra, Dr. J. A. Camarero

Department of Pharmacology and Pharmaceutical Sciences,
 University of Southern California
 Los Angeles, CA 90033 (USA)
 E-mail: jcamarero@usc.edu

Dr. J. A. Camarero
 Department of Chemistry, University of Southern California
 Los Angeles, CA 90033 (USA)

V. Lacey, Dr. L. Wang
 Jack H. Skirball Center for Chemical Biology and Proteomics,
 The Salk Institute for Biological Studies
 La Jolla, CA 92037 (USA)

S. Majumder, Dr. A. Shekhtman
 Department of Chemistry, State University of New York
 Albany, NY 12222 (USA)

[**] This work was supported by National Institutes of Health Research Grants R01-GM090323 (J.A.C.), R01-GM085006 (A.S.), DP2-OD004744 (L.W.) and by the Department of Defense Congressionally Directed Medical Research Program Grant PC09305 (J.A.C.). We would also like to thank Sam Ulin (Salk Institute) for his contribution in the preparation of plasmid pERAzi.

Supporting information for this article (experimental details) is available on the WWW under <http://dx.doi.org/10.1002/anie.201209219>.

of selenocysteine^[29] and pyrrolysine,^[30] which limits the chemical complexity of genetically-encoded libraries that can be produced in living cells. The recent development of nonsense suppressing orthogonal tRNA/synthetase technology has allowed the genetic encoding of a large variety of non-natural amino acids.^[31–34] Herein, we report the production of natively folded cyclotides containing non-natural amino acids by employing *in vivo* non-natural amino acid incorporation in combination with protein splicing to mediate intracellular backbone cyclization. To our knowledge, this is the first time that a natively folded cyclotide containing non-natural amino acids has been produced inside living cells. This approach makes possible *in vivo* generation of cyclotides containing non-natural amino acids with new or enhanced biological functions. For example, the ability to site-specifically introduce fluorescent probes should facilitate the *in vivo* production of fluorescently-labeled cyclotides for screening or probing molecular interactions using optical approaches.

To test the feasibility of introducing non-natural amino acids into folded cyclotides *in vivo*, we used the cyclotide MCoTI-I (Figure 1A). This cyclotide is a powerful trypsin inhibitor ($K_i \approx 20 \text{ pm}$)^[27] that has been recently isolated from dormant seeds of *Momordica cochinchinensis*, a plant member of the *Cucurbitaceae* family.^[35] Trypsin-inhibitor cyclotides are interesting candidates for drug design because their specificity can be altered and they can be used as natural scaffolds to generate novel binding activities.^[36]

Because MCoTI cyclotides have been expressed inside *Escherichia coli* cells using an intramolecular version of expressed protein ligation (EPL),^[27,28] we decided to try this approach first for *in vivo* generation of MCoTI-based cyclotides containing different non-natural amino acids. This method relies on the use of a protein splicing unit in combination with an *in vivo* intramolecular native chemical-ligation reaction to perform the backbone cyclization of the linear cyclotide precursor.^[26,27] The amber stop codon TAG was used to encode a non-natural amino acid at the position corresponding to the residue Asp¹⁴ in MCoTI-I. This residue is located in the middle of loop 2 (Figure 1A), which has been shown to be tolerant to mutations without affecting the structure and biological activity of the resulting cyclotide.^[28] The incorporation of non-natural amino acids into the cyclotide framework was tested using *p*-methoxyphenylalanine (OmeF) and *p*-azidophenylalanine (AziF), which have been incorporated into various recombinant proteins.^[33] More importantly, incorporation of AziF into the cyclotide framework should allow the site-specific incorporation of alkyne-containing fluorescent probes *in vivo*, using click chemistry.^[37]

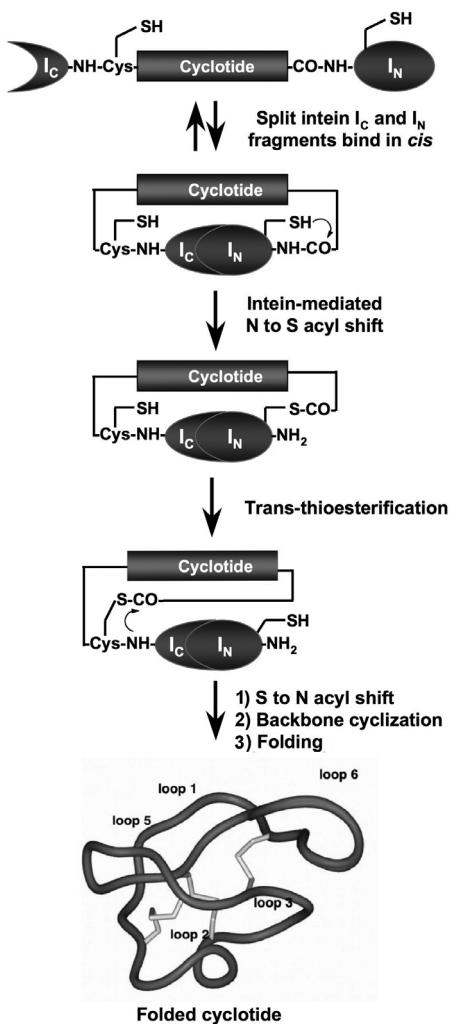
First, we explored the expression level of the corresponding intein precursors (**1b** and **1c**, Figure 1B) in BL21(DE3) cells. Expression of the intein precursors of MCoTI-OmeF and MCoTI-AziF was performed in cells co-transformed with a plasmid encoding the corresponding MCoTI-intein precursor for EPL-mediated cyclization and the plasmid encoding the orthogonal amber-suppressing tRNA_{CUA}/aminoacyl-tRNA synthetase pair specific for OmeF (pVLOmeRS) or AziF (pERAzi), respectively. In both cases, the expression level of the intein precursors (**1b** and **1c**) was similar (Supporting Information, Figure S1). The suppression effi-

ciency was estimated to be approximately 10% for **1b** and 20% for **1c**, compared to the expression of the wild-type MCoTI-I intein precursor **1a** (approximately 40 mg L⁻¹).

Under these expression conditions, all intein precursors showed around 60% *in vivo* cleavage, indicating that the intein was active and unaffected by the incorporation of the non-natural amino acid (Figure S1). Second, we tested the ability of the different intein-MCoTI precursors to produce the corresponding folded cyclotide by treatment with reduced glutathione (GSH) at pH 7.2 following the conditions optimized for MCoTI-cyclotides.^[38] In both cases, the *in vitro* reaction was clean and efficient in providing the properly folded cyclotides MCoTI-OmeF and MCoTI-AziF (Figure S1). The final yield after purification was 4 µg L⁻¹ (MCoTI-OmeF) and 14 µg L⁻¹ (MCoTI-AziF). The expression yield for the wild-type MCoTI-I using these expression and cyclization conditions was approximately 48 µg L⁻¹ after purification. Next, we explored the expression of the MCoTI-OmeF and MCoTI-AziF cyclotides inside bacterial cells using EPL-mediated cyclization.^[27,28,38] When we tried this approach with MCoTI-OmeF and MCoTI-AziF, however, the amount of folded cyclotides was below the detection limit.

To boost the expression of cyclotides *in vivo* we explored the use of protein trans-splicing (PTS) to facilitate the *in vivo* cyclization process and to improve the yield of non-natural amino acid-containing cyclotides (Scheme 1). PTS is a post-translational modification similar to protein splicing with the difference that the intein self-processing domain is split into N- (I_N) and C-intein (I_C) fragments. The split-intein fragments are not active individually, however, they can bind to each other with high specificity under appropriate conditions to form an active protein-splicing or intein domain in *trans*.^[39] PTS-mediated backbone cyclization can be accomplished by rearranging the order of the intein fragments. By fusing the I_N and I_C fragments to the C- and N-termini of the polypeptide for cyclization, the trans-splicing reaction yields a backbone-cyclized polypeptide (Scheme 1).^[40] This approach has recently been used for the biosynthesis of cyclic hexapeptides containing non-natural amino acids.^[41] Herein, *in vivo* cyclization was performed using the naturally occurring *Synechocystis* sp. (Ssp) PCC6803 DnaE split intein.^[42] However, the Ssp DnaE intein requires specific amino acid residues at both intein-extein junctions for efficient protein splicing.^[43] To overcome this problem we used the *Nostoc punctiforme* PCC73102 (Npu) DnaE split-intein. This DnaE intein has the highest reported rate of protein trans-splicing ($\tau_{1/2} \approx 60 \text{ s}$)^[44] and has a high splicing yield.^[44,45] We explored the ability of the Npu DnaE split-intein to produce folded wild-type MCoTI-I cyclotide in *E. coli* cells. To accomplish this, we designed the split-intein construct **2a** (Figure 1B). In this construct, the MCoTI-I linear precursor was fused in-frame at the C- and N-termini directly to the Npu DnaE I_N and I_C polypeptides. None of the additional native C- or N-extein residues were added in this construct. We used the native Cys residue located at the beginning of loop 6 of MCoTI-I (Figure 1A) to facilitate backbone cyclization. A histidine tag was also added at the N-terminus to facilitate purification.

In vivo expression of wild-type MCoTI-I using PTS-mediated backbone cyclization was achieved by transforming



Scheme 1. In vivo expression of native folded cyclotides using intein-mediated protein trans-splicing.

the plasmid encoding the split-precursor **2a** into Origami (DE3) cells to facilitate folding.^[46] The MCoTI-precursor split-intein was over expressed for 18 hours at room temperature, showing very high levels of expression (approximately 70 mg L^{-1}) and almost complete cleavage (more than 95% in vivo, Figure 2A). The high reactivity of this precursor prevented us from performing a full characterization. Next, we tried to isolate the natively folded MCoTI-I generated in vivo by incubating the soluble fraction of a fresh cell lysate with trypsin immobilized on sepharose beads. Correctly folded MCoTI-cyclotides are able to bind trypsin with high affinity ($K_i \approx 20\text{--}30 \text{ pm}$). Therefore, this step can be used for affinity purification and to test the biological activity of the recombinant cyclotides.^[28] After extensive washing, the absorbed products were eluted with a solution containing 8 M guanidinium chloride (GdmCl) and analyzed by HPLC. The HPLC analysis revealed the presence of a major peak that had the expected mass of the native MCoTI-I fold (Figure 2B; Figure S2). Recombinant MCoTI-I produced by PTS-mediated cyclization was also natively folded, as characterized by NMR spectroscopy (Figure S2 and Table S2).^[5]

The in vivo expression level of folded MCoTI-I produced by PTS-mediated cyclization was estimated to be 70 \mu g L^{-1} of bacterial culture, which corresponds to an intracellular concentration of approximately 7 \mu M .

In vivo expression of folded MCoTI cyclotides by PTS was about seven times more efficient than intramolecular EPL-mediated backbone cyclization. This improvement may be explained by our choice of the split-intein *Npu* DnaE. This split-intein is extremely efficient; it exhibits fast kinetics with a good yield of protein trans-splicing. Differences in the cyclization process between the PTS and EPL methods may also contribute to the improvement in the cyclization yield. In PTS, the cyclization is driven by the affinity between the two-intein fragments, *I_C* and *I_N*, which in the case of the *Npu* DnaE intein is very high ($K_D \approx 3 \text{ nm}$).^[47] Once the intein complex is formed, the trans-splicing reaction is also extremely fast ($\tau_{1/2} \approx 60 \text{ s}$ for the *Npu* DnaE intein).^[48] In contrast, EPL-mediated cyclization follows a slightly more complex mechanism that relies on the formation of the C-terminal thioester at the N-extein junction and the removal of the N-terminal leading sequence (a Met residue in this case) to provide an N-terminal Cys. These two groups then react to form a peptide bond between the N- and C-termini of the polypeptide. It is also worth noting that in contrast with the *Ssp* DnaE intein, which requires at least 4 native residues at the N- and C-terminal extein-intein junctions to work efficiently,^[49] the *Npu* orthologue used herein tolerates different sequences at both junctions as demonstrated by the efficient trans-splicing of precursor **2a** (Figure 2A). The tetrapeptide sequences at both intein-extein junctions in construct **2a** have only 20% sequence homology with the native sequences of both *Npu* DnaE exteins.

Encouraged by these results, we decided to try in vivo expression of cyclotides MCoTI-OmeF and MCoTI-AziF using PTS. For this purpose, precursors **2b** and **2c** (Figure 1B) were overexpressed in Origami (DE3) cells by transforming the plasmids pVLOmeRS and pERAzi and growing the bacterial cells in the presence of OmeF or AziF, respectively. Constructs **2b** and **2c** are similar to **2a** but were designed to incorporate non-natural amino acids instead of residue Asp¹⁴ in MCoTI-I (Figure 1B). The expression levels of the intein precursors **2b** and **2c** were approximately 7 mg L^{-1} (10% suppression in comparison to wild-type precursor **2a**) and approximately 20 mg mL^{-1} (25% suppression), respectively. In vivo trans-splicing for **2b** and **2c** was also similar ($\geq 95\%$, Figure 2A) to that of the wild-type PTS construct **2a**. Cyclotides MCoTI-OmeF and MCoTI-AziF were purified by affinity chromatography using trypsin sepharose beads from fresh soluble cell lysates, and the trypsin-bound fractions were analyzed by LC-MS/MS and ES-MS (Figure 2C; Figure S3). Cyclotide MCoTI-OmeF generated in vivo by PTS was also characterized by NMR, confirming the adoption of a native cyclotide fold (Figure 2D; Figure S4). The expression levels for cyclotides MCoTI-OmeF and MCoTI-AziF were estimated to be 1 \mu g L^{-1} and 2 \mu g L^{-1} , corresponding to an intracellular concentration of approximately 0.1 \mu M and 0.17 \mu M , respectively.

Next, we explored the possibility of using fluorescently-labeled cyclotides to perform screening of protein-cyclotide

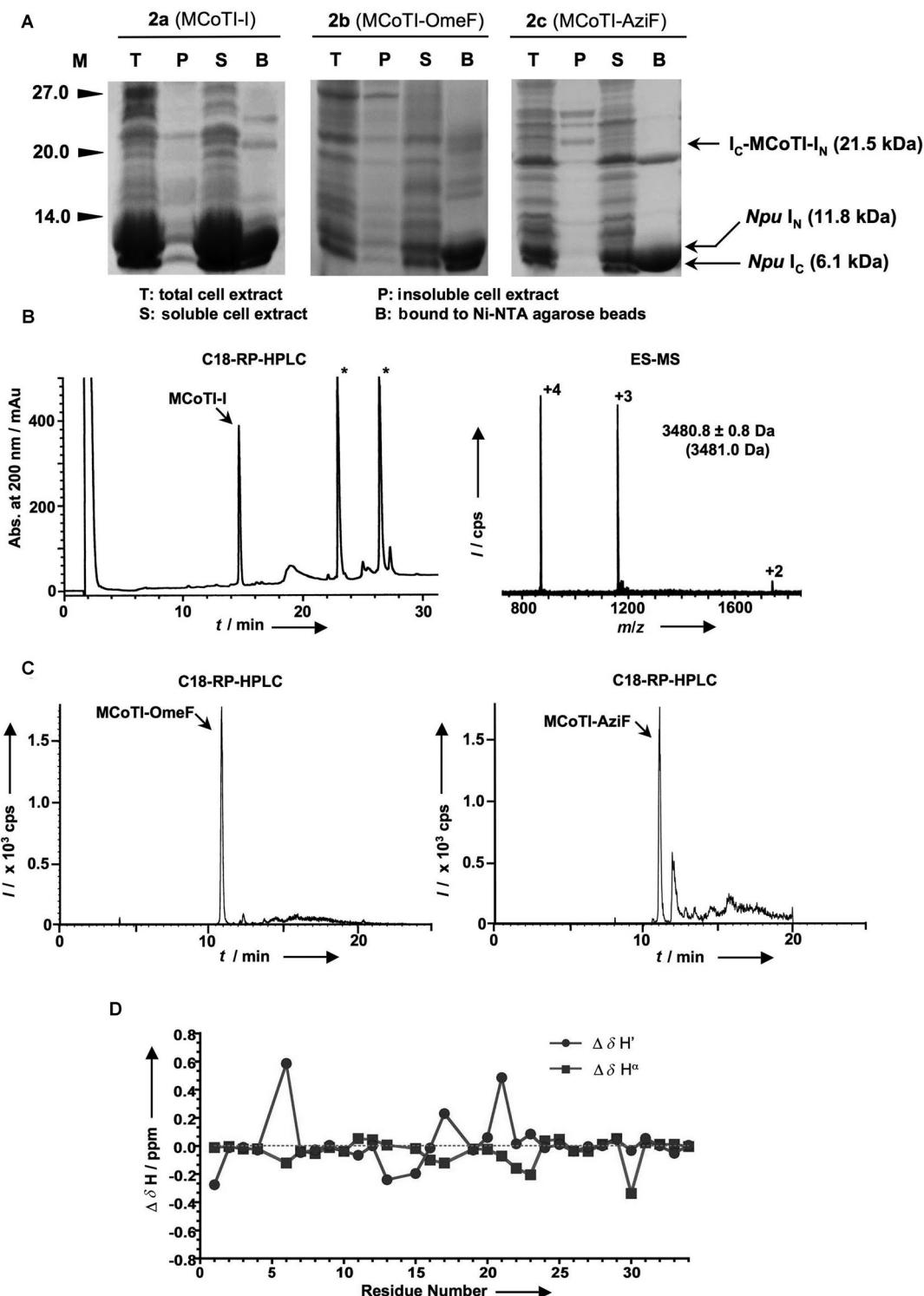


Figure 2. In vivo expression of MCoTi-I based cyclotides in *E. coli* cells using *Npu* DnaE intein-mediated PTS. A) SDS-PAGE analysis of the recombinant expression of cyclotide precursors **2a**, **2b**, and **2c** in Origami2(DE3) cells for production of the cyclotides MCoTi-I and MCoTi-OmeF, respectively. B) Analytical HPLC trace (left) of the soluble cell extract from bacterial cells expressing the precursor **2a** (MCoTi-I) after purification by affinity chromatography on a trypsin-sepharose column. Endogenous bacterial proteins that bind trypsin are marked with asterisks. Mass spectrum (right) of affinity purified MCoTi-I. The expected average molecular weight is shown in parentheses. C) LC-MS/MS analysis of the soluble cell extract of bacterial cells expressing precursors **2b** (MCoTi-OmeF) and **2c** (MCoTi-AziF). D) Summary of the backbone ^1H NMR assignments for the backbone protons of MCoTi-OmeF produced by PTS in vivo: $\Delta\delta(^1\text{H})$ are the deviations in the chemical shifts of the main chain protons between the values obtained for MCoTi-OmeF and MCoTi-I (Table S2). Assignments for residue 14 were not included in the graph.

interactions. To accomplish this, we used MCoTI-AziF and trypsin as a model system. Preliminary results showed that MCoTI-AziF can be efficiently (almost quantitatively) labeled in vitro with a dibenzylcyclooctyne (DBCO)-derivative of the fluorescent dye aminomethyl coumarin acetate (AMCA) through copper-free click chemistry (Figure S5). In vivo labeling of MCoTI-AziF with DBCO-AMCA was also very efficient (Figure 3A; Figure S7). No unreacted MCoTI-AziF was found after treatment of the cells with DBCO-AMCA, as determined by LC-MS/MS (data not shown). As expected, the resulting AMCA-labeled MCoTI-AziF was able to bind commercial porcine pancreatic trypsin efficiently (Figure S8),

indicating that introduction of the fluorophore in loop 2 of MCoTI-AziF did not have a detrimental effect on its biological activity. To monitor this interaction in vitro and in vivo, we used rat anionic trypsin, which can be expressed more efficiently in bacterial expression systems.^[50] We decided to use fluorescence resonance energy transfer (FRET) to visualize the interaction between trypsin and AMCA-labeled MCoTI-AziF. For this purpose, the protease was fused to the N-terminus of the enhanced green fluorescent protein (EGFP). AMCA and EGFP show a good overlap between the emission band of the donor (AMCA) and the absorption band of the acceptor (EGFP), thus allowing the visualization of their interaction using FRET.^[51] Moreover, structural analysis of a MCoTI-II–trypsin complex model^[28] reveals that the distance between the C-terminus of trypsin and the $\text{C}\alpha$ of residue 15 in MCoTI-I is approximately 35 Å. This distance is well within the range for the visualization of complex formation by FRET.^[51] The catalytic residue Ser¹⁹⁵ in trypsin was also mutated to Ala to facilitate the recombinant expression of trypsin-S195A-EGFP by preventing its cellular toxicity. As shown in Figure 3B,C, AMCA-labeled cyclotide MCoTI-AziF was able to efficiently bind trypsin-S195A-EGFP (K_D of 1.8 ± 0.7 nm) in vitro, and more importantly, the cyclotide–protein interaction could be easily monitored by intermolecular FRET shown by the simultaneous decrease and increase of the fluorescence signals at 445 nm and 515 nm, respectively.

In summary, we have shown that the biosynthesis of cyclotides containing non-natural amino acids can be achieved by using different intein-based methods. EPL-backbone cyclization can provide non-natural amino acid containing cyclotides when the cyclization is carried out in vitro by GSH-induced cyclization and folding of the corresponding precursor. In vivo production, however, is less efficient using this method. We have shown that PTS-mediated backbone cyclization using the highly efficient *Npu* DnaE split-intein can be used for the efficient production of cyclotides inside live *E. coli* cells. We estimate that the in vivo production of MCoTI-I was around seven times more efficient using *Npu* DnaE PTS than EPL, thereby providing an attractive alternative for the production of these types of polypeptides in bacterial cells. The high efficiency of PTS-mediated cyclization combined with nonsense-suppressing orthogonal tRNA/synthetase technology made the in vivo production of cyclotides containing non-natural amino acids possible. Of particular interest is the introduction of azido-containing amino acids, which can react with DBCO-containing fluorescent probes to provide in vivo fluorescently labeled cyclotides. The classical approach for in vivo production of fluorescent-labeled proteins by fusing a fluorescent protein to the target protein is not applicable to cyclotides, owing to their small size and restricted backbone-cyclized topology. We have shown that cyclotides containing the non-natural amino acid AziF can be expressed in live bacterial cells and easily labeled with DBCO-AMCA to monitor cyclotide–protein interactions. This finding opens the possibility for in vitro and potentially also in vivo screening of genetically encoded libraries of cyclotides for the rapid selection of novel cyclotide sequences able to bind a specific

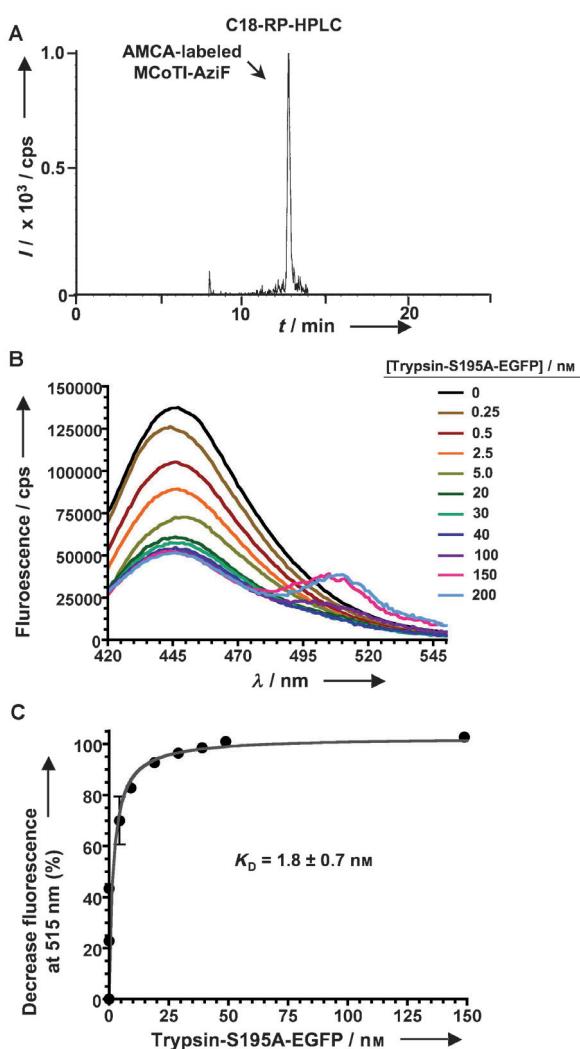


Figure 3. In vivo production and biological activity of AMCA-labeled MCoTI-AziF. A) LC-MS/MS trace of the soluble cell extract of bacterial cells indicating the presence of AMCA-labeled MCoTI-AziF. B) Titration of AMCA-labeled MCoTI-AziF (2 nm) with increasing amounts of trypsin-S195A-EGFP (0–200 nm) monitored by fluorescence spectroscopy. Excitation was performed at 360 nm. Background signal from EGFP in the absence of AMCA-labeled MCoTI-AziF was subtracted in all spectra. C) Binding isotherm of AMCA-labeled MCoTI-AziF and trypsin-S195A-EGFP by plotting the decrease in fluorescence signal at 445 nm. The dissociation constant (K_D) was calculated assuming a 1:1 molecular complex.

bait protein using high-throughput cell-based optical screening approaches.

Received: November 17, 2012

Published online: January 15, 2013

Keywords: circular proteins · click chemistry · cyclotides · non-natural amino acids · protein engineering

[1] D. J. Craik, S. Simonsen, N. L. Daly, *Curr. Opin. Drug Discov. Devel.* **2002**, *5*, 251.

[2] D. J. Craik, M. Cemazar, C. K. Wang, N. L. Daly, *Biopolymers* **2006**, *84*, 250.

[3] K. Jagadish, J. A. Camarero, *Biopolymers* **2010**, *94*, 611.

[4] A. E. Garcia, J. A. Camarero, *Curr. Mol. Pharmacol.* **2010**, *3*, 153.

[5] S. S. Puttamadappa, K. Jagadish, A. Shekhtman, J. A. Camarero, *Angew. Chem.* **2010**, *122*, 7184; *Angew. Chem. Int. Ed.* **2010**, *49*, 7030.

[6] N. L. Daly, K. J. Rosengren, D. J. Craik, *Adv. Drug Delivery Rev.* **2009**, *61*, 918.

[7] O. Saether, D. J. Craik, I. D. Campbell, K. Sletten, J. Juul, D. G. Norman, *Biochemistry* **1995**, *34*, 4147.

[8] K. P. Greenwood, N. L. Daly, D. L. Brown, J. L. Stow, D. J. Craik, *Int. J. Biochem. Cell Biol.* **2007**, *39*, 2252.

[9] J. Contreras, A. Y. Elnagar, S. F. Hamm-Alvarez, J. A. Camarero, *J. Controlled Release* **2011**, *155*, 134.

[10] L. Cascales, S. T. Henriques, M. C. Kerr, Y. H. Huang, M. J. Sweet, N. L. Daly, D. J. Craik, *J. Biol. Chem.* **2011**, *286*, 36932.

[11] R. J. Clark, N. L. Daly, D. J. Craik, *Biochem. J.* **2006**, *394*, 85.

[12] D. J. Craik, M. Cemazar, N. L. Daly, *Curr. Opin. Drug Discovery Dev.* **2006**, *9*, 251.

[13] D. J. Craik, N. L. Daly, J. Mulvenna, M. R. Plan, M. Trabi, *Curr. Protein Pept. Sci.* **2004**, *5*, 297.

[14] P. Thongyoo, N. Roque-Rosell, R. J. Leatherbarrow, E. W. Tate, *Org. Biomol. Chem.* **2008**, *6*, 1462.

[15] L. Cascales, D. J. Craik, *Org. Biomol. Chem.* **2010**, *8*, 5035.

[16] A. G. Poth, M. L. Colgrave, R. Philip, B. Kerenga, N. L. Daly, M. A. Anderson, D. J. Craik, *ACS Chem. Biol.* **2011**, *6*, 345.

[17] A. G. Poth, M. L. Colgrave, R. E. Lyons, N. L. Daly, D. J. Craik, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 10127.

[18] G. K. Nguyen, S. Zhang, N. T. Nguyen, P. Q. Nguyen, M. S. Chiu, A. Hardjojo, J. P. Tam, *J. Biol. Chem.* **2011**, *286*, 24275.

[19] J. P. Mulvenna, C. Wang, D. J. Craik, *Nucleic Acids Res.* **2006**, *34*, D192.

[20] C. K. Wang, Q. Kaas, L. Chiche, D. J. Craik, *Nucleic Acids Res.* **2008**, *36*, D206.

[21] J. P. Mulvenna, J. S. Mylne, R. Bharathi, R. A. Burton, N. J. Shirley, G. B. Fincher, M. A. Anderson, D. J. Craik, *Plant Cell* **2006**, *18*, 2134.

[22] C. W. Gruber, A. G. Elliott, D. C. Ireland, P. G. Delprete, S. Dessein, U. Goransson, M. Trabi, C. K. Wang, A. B. Kinghorn, E. Robbrecht, D. J. Craik, *Plant Cell* **2008**, *20*, 2471.

[23] N. L. Daly, S. Love, P. F. Alewood, D. J. Craik, *Biochemistry* **1999**, *38*, 10606.

[24] O. Avrutina, H. U. Schmoldt, H. Kolmar, U. Diederichsen, *Eur. J. Org. Chem.* **2004**, *4931*.

[25] P. Thongyoo, E. W. Tate, R. J. Leatherbarrow, *Chem. Commun.* **2006**, *2848*.

[26] R. H. Kimura, A. T. Tran, J. A. Camarero, *Angew. Chem.* **2006**, *118*, 987; *Angew. Chem. Int. Ed.* **2006**, *45*, 973.

[27] J. A. Camarero, R. H. Kimura, Y. H. Woo, A. Shekhtman, J. Cantor, *ChemBioChem* **2007**, *8*, 1363.

[28] J. Austin, W. Wang, S. Puttamadappa, A. Shekhtman, J. A. Camarero, *ChemBioChem* **2009**, *10*, 2663.

[29] A. Böck, K. Forchhammer, J. Heider, C. Baron, *Trends Biochem. Sci.* **1991**, *16*, 463.

[30] G. Srinivasan, C. M. James, J. A. Krzycki, *Science* **2002**, *296*, 1459.

[31] C. C. Liu, P. G. Schultz, *Annu. Rev. Biochem.* **2010**, *79*, 413.

[32] Q. Wang, A. R. Parrish, L. Wang, *Chem. Biol.* **2009**, *16*, 323.

[33] L. Wang, J. Xie, P. G. Schultz, *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 225.

[34] L. Wang, A. Brock, B. Herberich, P. G. Schultz, *Science* **2001**, *292*, 498.

[35] J. F. Hernandez, J. Gagnon, L. Chiche, T. M. Nguyen, J. P. Andrieu, A. Heitz, T. Trinh Hong, T. T. Pham, D. Le Nguyen, *Biochemistry* **2000**, *39*, 5722.

[36] A. Gould, Y. Ji, T. L. Aboye, J. A. Camarero, *Curr. Pharm. Des.* **2011**, *17*, 4294.

[37] K. E. Beatty, J. D. Fisk, B. P. Smart, Y. Y. Lu, J. Szychowski, M. J. Hangauer, J. M. Baskin, C. R. Bertozzi, D. A. Tirrell, *Chem. BioChem* **2010**, *11*, 2092.

[38] J. Austin, R. H. Kimura, Y. H. Woo, J. A. Camarero, *Amino Acids* **2010**, *38*, 1313.

[39] H. Sancheti, J. A. Camarero, *Adv. Drug Delivery Rev.* **2009**, *61*, 908.

[40] T. L. Aboye, J. A. Camarero, *J. Biol. Chem.* **2012**, *287*, 27026.

[41] T. S. Young, D. D. Young, I. Ahmad, J. M. Louis, S. J. Benkovic, P. G. Schultz, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 11052.

[42] C. P. Scott, E. Abel-Santos, M. Wall, D. Wahnon, S. J. Benkovic, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 13638.

[43] A. Tavassoli, S. J. Benkovic, *Nat. Protoc.* **2007**, *2*, 1126.

[44] J. Zettler, V. Schutz, H. D. Mootz, *FEBS Lett.* **2009**, *583*, 909.

[45] H. Iwai, S. Zuger, J. Jin, P. H. Tam, *FEBS Lett.* **2006**, *580*, 1853.

[46] P. H. Bessette, F. Aslund, J. Beckwith, G. Georgiou, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 13703.

[47] N. H. Shah, M. Vila-Perello, T. W. Muir, *Angew. Chem.* **2011**, *123*, 6641; *Angew. Chem. Int. Ed.* **2011**, *50*, 6511.

[48] R. Borra, D. Dong, A. Y. Elnagar, G. A. Woldemariam, J. A. Camarero, *J. Am. Chem. Soc.* **2012**, *134*, 6344.

[49] Y. Kwon, M. A. Coleman, J. A. Camarero, *Angew. Chem.* **2006**, *118*, 1758; *Angew. Chem. Int. Ed.* **2006**, *45*, 1726.

[50] J. R. Vasquez, L. B. Evnin, J. N. Higaki, C. S. Craik, *J. Cell. Biochem.* **1989**, *39*, 265.

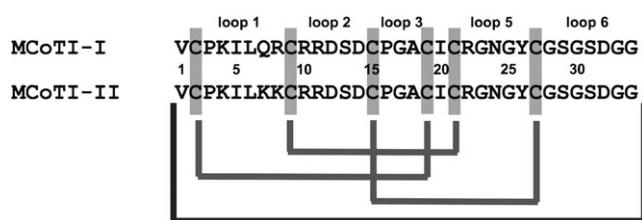
[51] S. M. Kuhn, M. Rubini, M. A. Muller, A. Skerra, *J. Am. Chem. Soc.* **2011**, *133*, 3708.

[52] M. E. Felizmenio-Quimio, N. L. Daly, D. J. Craik, *J. Biol. Chem.* **2001**, *276*, 22875.

Backbone Dynamics of Cyclotide MCoTI-I Free and Complexed with Trypsin**

Shadakshara S. Puttamadappa, Krishnappa Jagadish, Alexander Shekhtman, and Julio A. Camarero*

Cyclotides are a new emerging family of large plant-derived backbone-cyclized polypeptides (about 28–37 amino acids long) that share a disulfide-stabilized core (three disulfide bonds) characterized by an unusual knotted arrangement.^[1] Cyclotides contrast with other circular polypeptides in that they have a well-defined three-dimensional structure, and despite their small size can be considered as microproteins. Their unique circular backbone topology and knotted arrangement of three disulfide bonds makes them exceptionally stable to thermal and enzymatic degradation (Scheme 1).



Scheme 1. Primary structure and disulfide connectivities of MCoTI cyclotides. Dark gray and light gray connectors represent peptide and disulfide bonds, respectively.

Furthermore, their well-defined structures have been associated with a wide range of biological functions.^[2,3] Cyclotides MCoTI-I/II are powerful trypsin inhibitors ($K_i \approx 20$ –30 pm) that have been recently isolated from the dormant seeds of *Momordica cochinchinensis*, a plant member of the cucurbitaceae family.^[4] Although MCoTI cyclotides do not share significant sequence homology with other cyclotides beyond the presence of the three cystine bridges, structural analysis by NMR spectroscopy has shown that they adopt a similar

backbone-cyclic cystine-knot topology.^[5,6] MCoTI cyclotides, however, show high sequence homology with related cystine-knot squash trypsin inhibitors,^[4] and therefore represent interesting molecular scaffolds for drug design.^[7–10]

Determination of the backbone dynamics of these fascinating microproteins is key for understanding their physical and biological properties. Internal motions of a protein on different timescales, extending from picoseconds to a second, have been suggested to play an important role in its biological function.^[11] A better understanding of the backbone dynamics of the cyclotide scaffold will be extremely helpful for evaluating its utility as a scaffold for peptide-based drug discovery. Such insight will help in the design of optimal focused libraries that can be used for the discovery of new cyclotide sequences with novel biological activities.^[12,13]

Herein, we report for the first time the determination of the internal dynamics of the cyclotide MCoTI-I in the free state and complexed with trypsin. Uniformly ¹⁵N-labeled natively folded cyclotide MCoTI-I was recombinantly produced in *Escherichia coli* growing in minimal M9 medium containing ¹⁵NH₄Cl as the only source of nitrogen. Concomitant backbone cyclization and folding were accomplished by using intramolecular native chemical ligation^[14,15] in combination with a modified protein splicing unit (Figure S1, Supporting Information).^[16–18] The internal dynamics of cyclotide MCoTI-I was obtained from ¹⁵N spin-lattice and spin–spin relaxation times and ¹⁵N{¹H} heteronuclear Overhauser effect (NOE) enhancements.^[11] The backbone flexibility was characterized by the square of the generalized order parameter, S^2 , which reveals the dynamics of backbone NH groups on the pico- to nanosecond timescale.^[19,20] The order parameter satisfies the inequality $0 \leq S^2 \leq 1$, in which lower values indicate larger amplitudes of intramolecular motions. Motions on the milli- to microsecond timescale were assessed by the presence of the chemical exchange terms in the spin–spin relaxation.

The NMR spectrum and S^2 values, derived from the ¹⁵N relaxation data of free MCoTI-I, are shown in Figure 1a and d, respectively. Residues Ile5 and Gly23 of free MCoTI-I were excluded from the backbone dynamics analysis since the relaxation data could not be fitted to a monoexponential function, possibly as a result of chemical exchange.^[21] Gln7 was not assigned because of broadening of the NMR signal, presumably caused by fast exchange with water. The S^2 values for free MCoTI-I show that most of the NH groups of the cyclotide backbone are highly constrained with S^2 values ≥ 0.8 , thus resembling those found in well-folded globular proteins (Table 1). The average S^2 value, $\langle S^2 \rangle$, for free MCoTI-I was 0.83 ± 0.03 . This value is similar to that found

[*] Dr. K. Jagadish, Dr. J. A. Camarero
Department of Pharmaceutical Sciences and Pharmacology
University of Southern California
Los Angeles, CA 90033 (USA)
Fax: (+1) 323-224-7473
E-mail: jcamarero@usc.edu
S. S. Puttamadappa, Dr. A. Shekhtman
Department of Chemistry, State University of New York
Albany, NY 12222 (USA)

[**] This work was supported by funding from the School of Pharmacy at the University of Southern California, by National Institute of Health award GM090323-01 to J.A.C., and by American Diabetes Association award 1-06-CD-23 and National Institute of Health award GM090323-01 to A.S.

 Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/anie.201002906>.

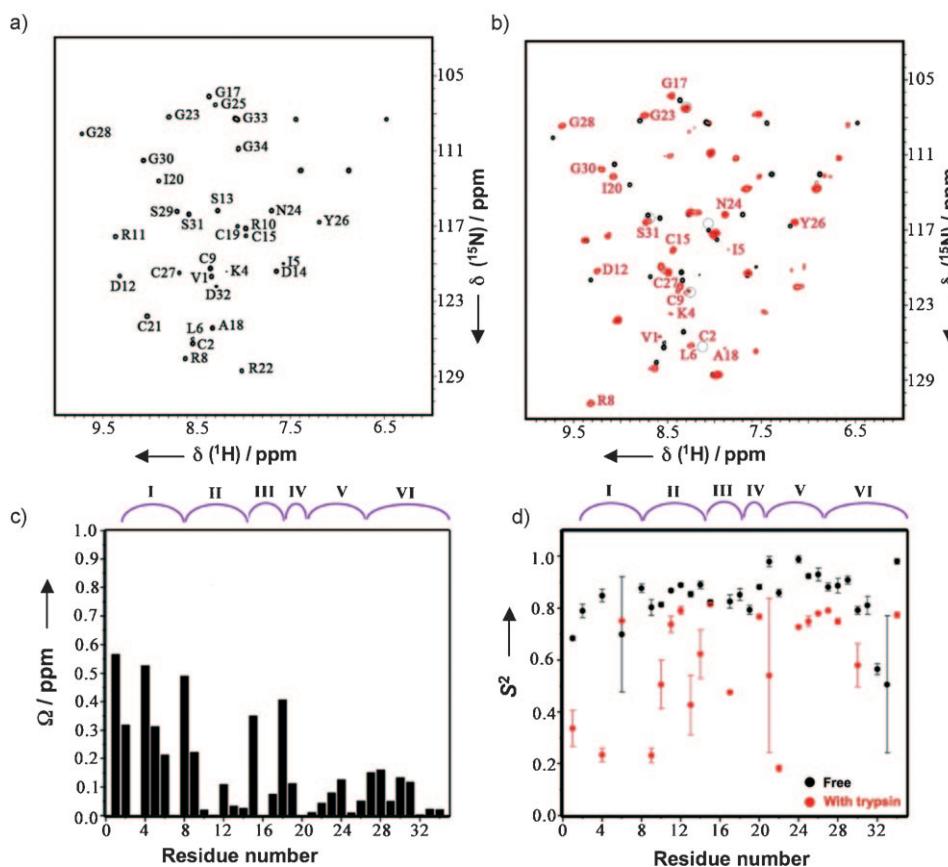


Figure 1. NMR analysis of the backbone dynamics of free and trypsin-bound MCoTI-I. a) $\{^{15}\text{N}, ^1\text{H}\}$ NMR heteronuclear single quantum correlation (HSQC) spectrum of free MCoTI-I. Chemical shift assignments of the backbone amides are indicated. b) Overlay of the $\{^{15}\text{N}, ^1\text{H}\}$ HSQC spectra of free (black) and trypsin-bound MCoTI-I (red). Residues with large average amide chemical shift differences between two different states ($>0.3 \text{ ppm}$) are indicated. Residues that are broadened in trypsin-bound MCoTI-I are indicated by gray circles. c) Average amide chemical shift difference for all the assigned residues in free and trypsin-bound MCoTI-I. The chemical shift difference was calculated as: $\Delta\Omega = [(\Delta\Omega_{\text{NH}}^2 + 0.04\Delta\Omega_{\text{N}}^2)/2]^{1/2}$, where $\Delta\Omega_{\text{NH}}$ and $\Delta\Omega_{\text{N}}$ are the changes in the amide proton and nitrogen chemical shifts (ppm), respectively. d) Order parameter, S^2 , for free (black) and trypsin-bound MCoTI-I (red). The S^2 value is a measure of backbone flexibility and represents the degree of angular restriction of the N-H vector in the molecular frame. The MCoTI-I loops are shown at the top of (c) and (d). Small unassigned peaks in the spectra of both free and trypsin-bound MCoTI-I are from a minor conformation of the protein, and result from a known isomerization of the backbone at an Asp-Gly sequence in loop 6 of MCoTI-I.

Table 1: Average order parameters of structural elements in MCoTI-I in the free state and bound to trypsin.

Structural element	Sequence	$\langle S^2 \rangle$	$\langle S^2 \rangle$
		Free MCoTI-I	Trypsin-MCoTI-I
loop 1	3–8	0.81 ± 0.01	0.49 ± 0.05
loop 2	10–14	0.81 ± 0.01	0.62 ± 0.07
loop 3	16–18	0.84 ± 0.02	$0.48^{[c]}$
loop 4	20	$0.88^{[c]}$	$0.76^{[c]}$
loop 5	22–26	0.92 ± 0.02	0.61 ± 0.01
loop 6	28–34	0.76 ± 0.05	0.61 ± 0.05
cystine knot	2, 10, 15, 19, 21, 27	0.84 ± 0.02	0.60 ± 0.08

[a] S^2 values for residues 5 and 23 from free MCoTI-I are not included in the average because the relaxation data could not be fitted to a monoexponential function. [b] S^2 values for residues 2, 5, 8, 18, 19, 23, 29, 31, 32, and 33 from trypsin-bound MCoTI-I are not included in the average because of the lack of signal intensity or because the relaxation data could not be fitted to a monoexponential function. [c] $\langle S^2 \rangle$ contains the S^2 value for a single residue.

for the six cystine residues involved in the cystine knot ($\langle S^2 \rangle = 0.84 \pm 0.02$) and is considerably larger than those found for other linear squash trypsin inhibitors ($\langle S^2 \rangle = 0.71$ for trypsin inhibitor from *Cucurbita maxima* (CMTI-III, 78% homology with MCoTI-I)),^[22] thus indicating the importance of the backbone cyclization to rigidifying the overall structure. Loops 2 through 5 in free MCoTI-I showed $\langle S^2 \rangle$ values ≥ 0.8 . In particular, loop 5 showed an $\langle S^2 \rangle$ value of 0.92 ± 0.02 , well above the average for the molecule and the cystine knot. In contrast, loops 1 and 6 showed $\langle S^2 \rangle$ values below the average for the molecule. Thus, loop 6, which is believed to act as a very flexible linker to allow cyclization,^[23] had an $\langle S^2 \rangle$ value of 0.76 ± 0.17 with only two residues, Asp32 and Gly33, having values below 0.6. Despite this small $\langle S^2 \rangle$ value, residues in loop 6 did not require significant chemical exchange terms (Figure S2 and Table S1, Supporting Information), which suggests that the mobility observed arises mostly from local vibrations.

The $\langle S^2 \rangle$ value for loop 1, which is responsible for binding trypsin, was 0.81 ± 0.07 . This value is $\approx 90\%$ of the average value for free MCoTI-I. Residue Leu6 in

loop 1 also required chemical exchange terms to be considered, thus indicating the existence of intramolecular conformational exchange on the micro- to millisecond timescale. The mobility observed in loop 1 at both nano- to picosecond and millisecond timescales has also been described in other trypsin inhibitors,^[22, 24, 25] and it has been suggested to play an important role in receptor-ligand binding.^[11]

To explore whether that was the case in the MCoTI cyclotides, we next studied the effect of ligand binding on the backbone dynamics of MCoTI-I (Figures 1 and 2). To exclude the possibility that trypsin could cleave or scramble the disulfide bonds of MCoTI-I upon complex formation, we used a competition experiment of trypsin- $[^{15}\text{N}]$ MCoTI-I with unlabeled MCoTI-I. The results indicated that the structure of MCoTI-I is unaltered upon trypsin binding (Figure S3, Supporting Information). Trypsin binding led to large ($>0.3 \text{ ppm}$) and specific changes in the chemical shifts of

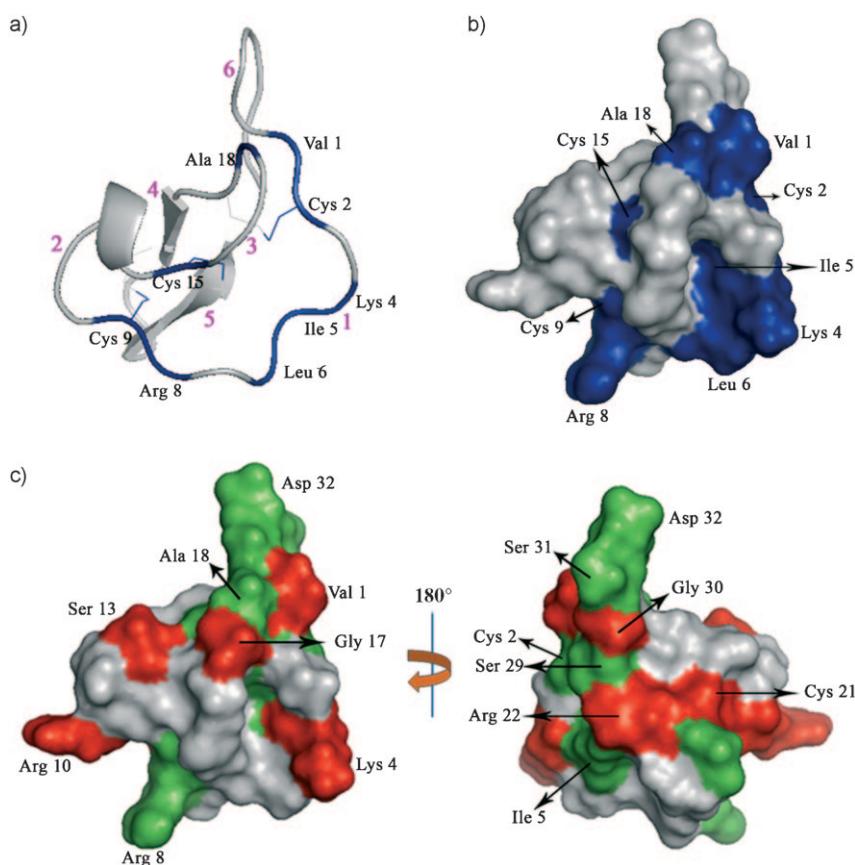


Figure 2. Trypsin binding to MCoTI-I affects the MCoTI-I backbone dynamics. a) Ribbon and b) surface diagrams of the trypsin–MCoTI-I interaction map. Red numbers indicate the positions of the MCoTI-I loops. The MCoTI-I residues with a large chemical shift difference (>0.3 ppm) are in blue. c) Changes in the MCoTI-I order parameter as a result of binding to trypsin. Residues with $S_f^2 - S_b^2 > 0.2$, where S_f^2 and S_b^2 are the order parameters of the free and trypsin-bound MCoTI-I, respectively, are depicted in red. MCoTI-I residues that were broadened in $^{15}\text{N}, ^1\text{H}$ HSQC because of binding to trypsin are shown in green. The structure of free MCoTI-II (PDB code: 1IB9)^[6] was used to illustrate the changes of MCoTI-I dynamics arising from trypsin binding.

the residues located in loop 1 (Cys2, Lys4, Ile5, Arg8), loop 3 (Cys15 and Ala18), and loop 6 (Val1) (Figures 1c and 2b and Table S2, Supporting Information). NMR signals of Cys2, Ile5, Cys19, Ser29, Ser31, Asp32, and Gly33 were significantly broadened, presumably because of intramolecular chemical exchanges in the trypsin–MCoTI-I complex. Arg8, Ala18, and Gly23 were excluded from backbone dynamics analysis because their peaks were broadened in the $^{15}\text{N}\{^1\text{H}\}$ NOE spectra. Similar findings have already been reported for other biomolecular interactions.^[26]

We used these changes to construct the trypsin–MCoTI-I interaction surface. The binding surface is contiguous and spans 46% of the total molecular area of MCoTI-I (Figure 2b). As expected, the major difference in the backbone dynamics was observed in the binding loop (Table 1), where the mobility in the nano- to picosecond timescale was increased in MCoTI once bound to trypsin. Loop 1 showed $\langle S^2 \rangle = 0.49 \pm 0.02$, which is much lower than the value for the rest of the molecule ($\langle S^2 \rangle = 0.65 \pm 0.07$). Several residues in loop 2 (Cys9, Arg10, Ser13, and Asp14), loop 3 (Gly17), and loop 5 (Cys21 and Arg22) also showed signifi-

cantly lower values of S^2 upon complex formation (Figures 1d and 2c). It is likely that the increase in mobility observed in these loops may help to accommodate the increased flexibility of the binding loop (Figure 2c).

Since our data clearly show that backbone flexibility of cyclotide MCoTI-I increases significantly upon binding to trypsin, we decided to estimate the contribution of these motions to the overall Gibbs free energy of binding (ΔG). The energetic benefit of this increase in backbone flexibility can be estimated from the experimental relaxation data, by using the experimentally measured order parameters S^2 .^[27] The estimated ΔG value was approximately -62 kJ mol^{-1} at 298 K. This value is almost identical to the calculated value from the trypsin inhibitory constant of MCoTI-I ($K_i \approx 20 \text{ pm}$,^[28] $\Delta G \approx -61 \text{ kJ mol}^{-1}$). The calculated entropic contribution ($-T\Delta S$) at the same temperature was approximately -46 kJ mol^{-1} . These results highlight the importance of the backbone entropic term to the formation of the trypsin–MCoTI-I complex, although a more detailed thermodynamic analysis that also includes the side-chain motions may be required.

In summary, we have reported the backbone dynamics of the cyclotide MCoTI-I in the free state and complexed to its binding partner trypsin in solution. To our knowledge this is the first time the backbone dynamics of a natively folded cyclotide has been reported. This has

been possible because of the use of modified protein splicing units for the heterologous expression of folded cyclotides using bacterial expression systems^[17, 18, 29] to incorporate NMR-active nuclei such as ^{15}N . Our results on the backbone dynamics of free cyclotide MCoTI-I confirm that MCoTI-I adopts a well-folded and highly compact structure with an $\langle S^2 \rangle$ value of 0.83. This value is similar to those found in the regions of well-folded proteins with restricted backbone dynamics.

The results also indicate that the trypsin-binding loop (loop 1) has a smaller S^2 value than the average value for the whole molecule, thus indicating a higher mobility of this region in the pico- to nanosecond timescale. This region also showed significant conformational exchange motions in the micro- to millisecond timescale. Loop 6 also possesses a higher mobility in the pico- to nanosecond timescale than the averaged value for MCoTI-I, although no significant conformational exchange motions were detected in the micro- to millisecond timescale. This result is intriguing since this loop contains a potentially flexible Gly–Ser-rich sequence that is mostly absent among other linear trypsin squash inhibitors,

and therefore it was thought to be a highly flexible linker to allow cyclization. More surprising, however, was the fact that the backbone of MCoTI-I, and especially loop 1, increased the pico- to nanosecond mobility when bound to trypsin. This interesting result has already been observed in other high-affinity protein–protein interactions.^[30,31]

The thermodynamic analysis of the backbone contribution to the formation of the trypsin–MCoTI-I complex by using measured S^2 values also revealed the importance of the backbone entropic term in the formation of the complex. Similar findings have also been found in other protease inhibitors.^[32] This increment in backbone mobility may help to minimize the entropic penalties required for binding. Hence, we also observed in the HSQC spectrum of the trypsin–MCoTI-I complex the appearance of a signal corresponding to the $\epsilon\text{-NH}_3^+$ of Lys4 located in loop 1, which suggests that the ammonium group is protected and more rigid when forming the complex (Figure 3). This is the only

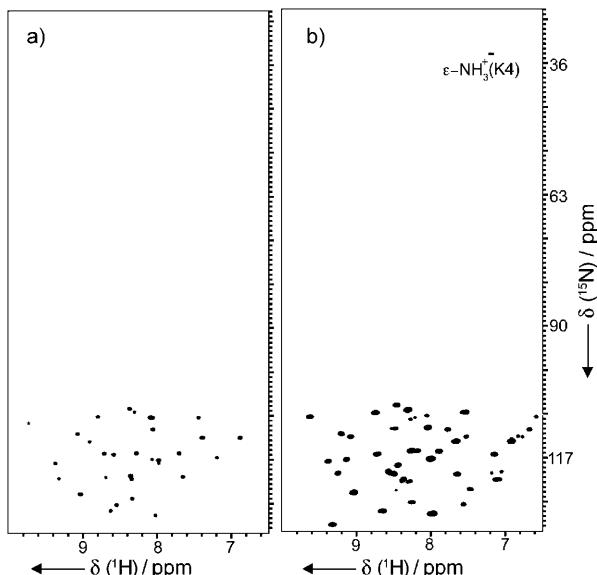


Figure 3. $\epsilon\text{-NH}_3^+$ of Lys4 is protected from fast exchange with the solvent in trypsin-bound MCoTI-I. $\{^{15}\text{N}, ^1\text{H}\}$ -HSQC spectra of free (a) and trypsin-bound MCoTI-I (b) were collected at room temperature with the ^{15}N -carrier position at 82 ppm and ^{15}N radio-frequency field strengths of 5.2 kHz for 90° and 180° pulses and 1.2 kHz for composite decoupling during acquisition.

Lys residue present in the sequence of MCoTI-I (Scheme 1) and therefore it can be unambiguously assigned. This residue is key for binding to trypsin^[29] and is responsible for binding to the specificity pocket of trypsin. This cross-peak was totally absent in the free MCoTI-I sample, which indicates that the $\epsilon\text{-NH}_3^+$ of Lys4 is less rigid and rapidly exchanging with solvent (Figure 3a).

Similarly, the broadening of aliphatic resonances for Arg side chains with essentially rigid guanidinium groups (that is, $^e\text{N}-\text{H}$ bond vectors) has also been described for protein–peptide complexes.^[26] Palmer and co-workers have recently suggested that this dynamic decoupling between the side-chain terminus from the rest of the aliphatic part of the side

chain may be a general biophysical strategy for maximizing residual side-chain and potentially backbone conformational entropy in proteins and their complexes,^[33] which is in agreement with our observations regarding the increase in MCoTI-I backbone mobility upon complex formation.

We have also mapped the binding surface of MCoTI-I once bound to trypsin. Major changes in chemical shifts were observed for the solvent-exposed residues located in loops 1 (Lys4, Ile5, Arg8), 3 (Ala18), and 6 (Val1) (Figures 1c and 2b). In agreement with these results, we have recently shown that the introduction of nonconservative mutations in these positions has a negative effect on the affinity for trypsin,^[29] thus indicating that they may be in close contact with the protease at the binding interface of the molecular complex.

Cyclotides present several characteristics that make them appear as promising leads or frameworks for peptide drug design.^[7,8] Investigation of the backbone dynamics is crucial for a better understanding of the dynamic structural properties of the cyclotide scaffold and how it affects the mode of binding of these interesting molecules. The reported data will help in the design of cyclotide-based libraries for molecular screening and the selection of *de novo* sequences with new biological activities, or the development of grafted analogues for use as peptide-based drugs.^[9,10]

Received: May 14, 2010

Published online: August 16, 2010

Keywords: cyclotides · molecular dynamics · NMR spectroscopy · protein–protein interactions · structural biology

- [1] N. L. Daly, K. J. Rosengren, D. J. Craik, *Adv. Drug Delivery Rev.* **2009**, *61*, 918.
- [2] D. J. Craik, S. Simonsen, N. L. Daly, *Curr. Opin. Drug Discovery Dev.* **2002**, *5*, 251.
- [3] D. J. Craik, M. Cemazar, C. K. Wang, N. L. Daly, *Biopolymers* **2006**, *84*, 250.
- [4] J. F. Hernandez, J. Gagnon, L. Chiche, T. M. Nguyen, J. P. Andrieu, A. Heitz, T. Trinh Hong, T. T. Pham, D. Le Nguyen, *Biochemistry* **2000**, *39*, 5722.
- [5] A. Heitz, J. F. Hernandez, J. Gagnon, T. Trinh Hong, T. T. Pham, T. M. Nguyen, D. Le-Nguyen, L. Chiche, *Biochemistry* **2001**, *40*, 7973.
- [6] M. E. Felizmenio-Quimio, N. L. Daly, D. J. Craik, *J. Biol. Chem.* **2001**, *276*, 22875.
- [7] R. J. Clark, N. L. Daly, D. J. Craik, *Biochem. J.* **2006**, *394*, 85.
- [8] D. J. Craik, M. Cemazar, N. L. Daly, *Curr. Opin. Drug Discovery Dev.* **2006**, *9*, 251.
- [9] D. J. Craik, N. L. Daly, J. Mulvenna, M. R. Plan, M. Trabi, *Curr. Protein Pept. Sci.* **2004**, *5*, 297.
- [10] P. Thongyoo, N. Roque-Rosell, R. J. Leatherbarrow, E. W. Tate, *Org. Biomol. Chem.* **2008**, *6*, 1462.
- [11] V. A. Jarymowycz, M. J. Stone, *Chem. Rev.* **2006**, *106*, 1624.
- [12] H. Sancheti, J. A. Camarero, *Adv. Drug Delivery Rev.* **2009**, *61*, 908.
- [13] D. C. Ireland, M. L. Colgrave, N. L. Daly, D. J. Craik, *Adv. Exp. Med. Biol.* **2009**, *611*, 477.
- [14] J. A. Camarero, T. W. Muir, *Chem. Commun.* **1997**, 1369.
- [15] J. A. Camarero, J. Pavel, T. W. Muir, *Angew. Chem.* **1998**, *110*, 361; *Angew. Chem. Int. Ed.* **1998**, *37*, 347.
- [16] J. A. Camarero, T. W. Muir, *J. Am. Chem. Soc.* **1999**, *121*, 5597.

[17] R. H. Kimura, A. T. Tran, J. A. Camarero, *Angew. Chem.* **2006**, *118*, 987; *Angew. Chem. Int. Ed.* **2006**, *45*, 973.

[18] J. A. Camarero, R. H. Kimura, Y. H. Woo, A. Shekhtman, J. Cantor, *ChemBioChem* **2007**, *8*, 1363.

[19] G. Lipari, A. Szabo, *J. Am. Chem. Soc.* **1982**, *104*, 4546.

[20] G. Lipari, A. Szabo, *J. Am. Chem. Soc.* **1982**, *104*, 4559.

[21] A. Horská, J. Horsky, R. G. S. Spencer, *J. Magn. Reson. Ser. A* **1994**, *110*, 82.

[22] J. Liu, Y. Gong, O. Prakash, L. Wen, I. Lee, J. K. Huang, R. Krishnamoorthi, *Protein Sci.* **1998**, *7*, 132.

[23] A. Heitz, O. Avrutina, D. Le-Nguyen, U. Diederichsen, J. F. Hernandez, J. Gracy, H. Kolmar, L. Chiche, *BMC Struct. Biol.* **2008**, *8*, 54.

[24] J. Liu, O. Prakash, M. Cai, Y. Gong, Y. Huang, L. Wen, J. J. Wen, J. K. Huang, R. Krishnamoorthi, *Biochemistry* **1996**, *35*, 1516.

[25] B. Szenthe, Z. Gaspari, A. Nagy, A. Perczel, L. Graf, *Biochemistry* **2004**, *43*, 3376.

[26] S. M. Pascal, T. Yamazaki, A. U. Singer, L. E. Kay, J. D. Forman-Kay, *Biochemistry* **1995**, *34*, 11353.

[27] A. G. Palmer III, *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 129.

[28] O. Avrutina, H. U. Schmoldt, D. Gabrijelcic-Geiger, D. Le Nguyen, C. P. Sommerhoff, U. Diederichsen, H. Kolmar, *Biol. Chem.* **2005**, *386*, 1301.

[29] J. Austin, W. Wang, S. Puttamadappa, A. Shekhtman, J. A. Camarero, *ChemBioChem* **2009**, *10*, 2663.

[30] K. Dutta, H. Shi, E. R. Cruz-Chu, K. Kami, R. Ghose, *Biochemistry* **2004**, *43*, 8094.

[31] R. Ghose, A. Shekhtman, M. J. Goger, H. Ji, D. Cowburn, *Nat. Struct. Biol.* **2001**, *8*, 998.

[32] S. Arumugam, G. Gao, B. L. Patton, V. Semenchenko, K. Brew, S. R. Van Doren, *J. Mol. Biol.* **2003**, *327*, 719.

[33] N. Trbovic, J. H. Cho, R. Abel, R. A. Friesner, M. Rance, A. G. Palmer III, *J. Am. Chem. Soc.* **2009**, *131*, 615.



NIH Public Access

Author Manuscript

Curr Mol Pharmacol. Author manuscript; available in PMC 2012 April 17.

Published in final edited form as:
Curr Mol Pharmacol. 2010 November 1; 3(3): 153–163.

Biological Activities of Natural and Engineered Cyclotides, a Novel Molecular Scaffold for Peptide-Based Therapeutics

Angie E. Garcia and Julio A. Camarero*

Department of Pharmacology and Pharmaceutical Sciences, School of Pharmacy, University of Southern California, Los Angeles, CA 90033, USA

Abstract

Cyclotides are a growing family of large plant-derived backbone-cyclized polypeptides (≈ 30 amino acids long) that share a disulfide-stabilized core characterized by an unusual knotted structure. Their unique circular backbone topology and knotted arrangement of three disulfide bonds makes them exceptionally stable to thermal, chemical, and enzymatic degradation compared to other peptides of similar size. Currently more than 100 sequences of different cyclotides have been characterized and the number is expected to increase dramatically in the coming years. Considering their stability, biological activities and ability to cross the cell membrane, cyclotides can be exploited to develop new peptide-based drugs with high potential for success. The cyclotide scaffold can be engineered or evolved using molecular evolution to inhibit protein-protein interactions implicated in cancer and other human diseases, or design new antimicrobial. The present review reports the biological diversity and therapeutic potential of natural and engineered cyclotides.

Keywords

Cyclic peptide; Cyclotides; Intein Kalata B1; MCoTI-I/II; Protein splicing; Protein engineering

Introduction

Head-to-tail or backbone-cyclized peptides are present throughout nature from bacteria to animals. Bacteria and fungi express numerous backbone-cyclized peptides that are currently in use as therapeutic agents [1]. For example, cyclosporin A is a fungal peptide with potent immunosuppressive properties and is used to treat organ transplant patients [2]. Daptomycin, is a 13-amino acid cyclic lipopeptide with a decanoyl side chain isolated from *Streptomyces roseosporus* that has recently been approved to treat infections against Gram-positive organisms, including multi-resistant strains [3]. In animals, the only known circular peptides are θ -defensins, which are expressed in blood leukocytes and bone marrow of Old World monkeys [4, 5]. θ -Defensins are antimicrobial peptides with broad-spectrum activities against bacteria, fungi, and viruses [6–8]. Backbone cyclized peptides have been also found in plants [9]. Sunflower trypsin inhibitor 1 (SFTI-1) for example is a bicyclic 14-residue long peptide found in sunflower seeds. SFTI-1 is the most potent known naturally occurring Bowman-Birk trypsin inhibitor [10]. Cyclotides, a novel family of small globular backbone-cyclized micro-proteins (≈ 30 -residues long), are also naturally found in plants. Here, we review the properties of cyclotides and the latest developments in the use of the cyclotide scaffold to design novel peptide-based therapeutics.

*Address all correspondence to: Associate Professor Department of Pharmacology and Pharmaceutical Sciences University of Southern California 1985 Zonal Avenue, PSC 616 Los Angeles, CA 90033 jcamare@usc.edu.

Cyclotides, a novel ultrastable molecular scaffold

Cyclotides are small globular micro-proteins with a unique head-to-tail cyclized backbone, which is stabilized by three disulfide bonds (Fig. 1). Currently, over 140 sequences have been identified in the plant species *Rubiaceae*, *Violaceae*, and *Cucurbitaceae* [11]. Natural cyclotides have various activities including insecticidal [12, 13], uterotonic [14], anti-HIV [15], antimicrobial [16, 17], antitumor [18], antihelminthic [19, 20] and have been reported to cross cell membranes [21]. Their insecticidal and antihelminthic properties suggest that they may function as defense molecules in plants.

Cyclotides share a unique head-to-tail circular knotted topology of three disulfide bridges, with one disulfide penetrating through a macrocycle formed by the two other disulfides and inter-connecting peptide backbones, forming what is called a cyclic cystine knot (CCK) motif (Fig. 1). The CCK topology is responsible for the high stability of cyclotides to enzymatic, thermal and chemical degradation [22]. The cyclotide family is divided into three structurally distinct families, Möbius, bracelet, and trypsin inhibitor subfamilies (Fig. 1). Möbius cyclotides are distinguished from bracelet cyclotides by the presence of a *cis*-Pro residue in loop 5. Trypsin inhibitor cyclotides have very different primary structures from Möbius and bracelet cyclotides, but retain the conserved cystine knot motif. Trypsin inhibitor cyclotides share a high sequence homology with related cystine-knot trypsin inhibitors found in squash such as EETI-II (*Ecballium elaterium* trypsin inhibitor II), and in fact can be considered cyclized homologs of these protease inhibitors. Thus, cyclotides can be considered natural combinatorial peptide libraries structurally constrained by the cystine-knot scaffold [23] and head-to-tail cyclization but are permissive of hypermutation of essentially all residues with the exception of the strictly conserved cysteines that comprise the knot [24-27].

Hence, cyclotides form a unique family of structurally-related peptides that possess remarkable stability due to the cystine knot, a small size making them readily accessible to chemical synthesis, and an excellent tolerance to sequence variations. Moreover, the first cyclotide to be discovered, kalata B1, is an orally effective uterotonic [14]. Intriguingly, the cyclotide MCoTI-II has also been shown to cross cell membranes through macropinocytosis [21]. We have also recently found that MCoTI-I has similar cellular-uptake properties to MCoTI-II (unpublished results). All of these features make cyclotides ideal tools for the development of a total novel class of peptide-based therapeutics.

Cyclotide biosynthesis

Cyclotides are ribosomally synthesized as precursor proteins, which consist of an endoplasmic reticulum (ER)-targeting sequence, a pro-region, a highly conserved N-terminal repeat (NTR) region, a mature cyclotide domain, and a C-terminal tail (Fig. 2). The combined NTR-cyclotide segment may contain one copy of the cyclotide sequence or there may be multiple copies of the same or different cyclotide sequences separated by additional NTR sequences. The precursor undergoes post-translational processing to generate a circular peptide by a mechanism that has not been completely elucidated yet [28, 29]. It has been hypothesized that a conserved Asn (or Asp) at the C-terminal cleavage site may be a recognition site by asparaginyl endoprotease (AEP) for cyclization of the peptide in plants [28, 29]. AEP has been shown to be involved in the post-translational processing of concanavalin A from the jackbean [30], and therefore, it is possible this enzyme may be involved in the cyclization of other plant peptides. Studies using transgenic plants that express a cyclotide precursor have demonstrated the involvement of AEP and requirement for the asparagine residue in the cyclotide sequence [28, 29]. The authors showed that inhibition of AEP led to a decrease in the amount of cyclic product and an accumulation of

linear peptides that were transiently expressed [29]. In addition, a complementary study showed that mutation of the asparagine residue or truncation of the conserved C-terminal tripeptide in transgenic plants resulted in no circular peptide production [28].

Biological activities of naturally occurring cyclotides

The natural function of cyclotides appears to be in protection of plants against insects [12, 13], nematodes [20, 25], and mollusks [31]. Studies have demonstrated that cyclotides can suppress the growth and development of insect and nematode larvae. Various other studies have also shown cyclotides have antimicrobial, hemolytic, uterotonic, and anti-HIV activities. Much of these activities likely involve interaction of the cyclotide with membranes, although the mechanism of action is not totally well understood.

The first cyclotide discovered, kalata B1, was identified in the plant *Oldenlandia affinis* in central Africa in the 1960's [32]. This plant was used by the natives to make a tea extract that was used to accelerate childbirth during labor [14, 33]. The main active ingredient in the tea extract was found to be a peptide that was named kalata B1, after the local name for the native medicine. The uterotonic properties of kalata B1 indicated that the peptide was orally bioavailable. The complete sequence, Cys-knotted arrangement, and cyclic nature of kalata B1 was determined 25 years after its original discovery [34]. Since the discovery of kalata B1, many more related cyclotides have been discovered and found to have various biological activities [35, 36].

Cyclotides were initially hypothesized to have antimicrobial activities based on the presence of hydrophilic and hydrophobic patches, which give an amphipathic character similar to classical antimicrobial peptides. The antimicrobial activities of cyclotides have been reported by two groups with conflicting results on the potency of kalata B1 against *Escherichia coli* and *Staphylococcus aureus*. In one study, kalata B1 was active against *S. aureus*, but not *E. coli* [16], and in the second study, the peptide had the reverse effect [17]. This is likely due to the technical differences in the experiments. Although kalata cyclotides are amphipathic, the overall charge is close to zero at neutral pH, making it unlikely that they interact with bacterial membranes electrostatically similar to classical cationic antimicrobial peptides [37]. Further studies are necessary to investigate the mechanisms of antimicrobial action given the growing occurrence of antibiotic resistance by microorganisms.

The anti-HIV properties of cyclotides have been extensively studied [15, 38-40]. They appear to mainly act by inhibiting viral entry into host cells as studies have shown a dose-dependent increase in cytoprotection [15]. This suggests the peptides may block binding or fusion of the virus, but the mechanism remains unclear. In general, cyclotide bioactivities appear to involve interactions with membranes and, therefore, this may be a mechanism for anti-HIV activities, by preventing fusion of the viral and host cell membranes. Studies have shown cyclotides can bind to model lipid membranes by surface plasmon resonance [41], and that binding occurs mainly through the peptide hydrophobic patches exposed on the surface [42-44]. This suggests membrane binding may be one mode for cyclotide activity against microorganisms.

Cyclotide interactions with membranes have also been suggested as their mechanism for cytotoxic activity. Studies have demonstrated antitumor activities of cyclotides, which were selective against cancer cell lines and solid tumors compared to normal mammalian cells [18, 45, 46]. Cancer cells differ from normal cells in the lipid and glycoprotein composition, which alters the overall net charge. The different potencies between cyclotide cytotoxicity are related to the three-dimensional structure as well as specific amino acid residues within the sequence [46, 47].

In addition to having antimicrobial and antitumor activities, some cyclotides have been found to cause extensive hemolysis of human and rat erythrocytes [16, 40, 48]. The cyclotide kalata B1 has strong hemolytic activity, although this can be eliminated by mutation to Ala of any one of eight residues located in the bioactive face of the molecule [25]. A more recent study has also shown that the hemolytic activity of kalata B1 could also be reduced or completely eliminated by mutation to Lys of any of the residues involved in either the bioactive or hydrophobic faces of kalata B1 [49]. The same study showed that the hemolytic, insecticidal and nematocidal activities of the different kalata B1 mutants were correlated, indicating there may be a common mechanism involving a cyclotide-membrane interaction [49]. The same authors have also shown recently the all D-analogue of kalata had similar nematicidal activity than the native peptide, thus corroborating that the biological activity is mediated by a cellular receptor [20]. In contrast, other cyclotides such as MCOTI-cyclotides have shown little or no hemolytic activity, thus demonstrating the diversity of cyclotide properties.

Another cyclotide with interesting biological activity is cyclopsychotride (Cpt) A. Cpt A is natural cyclotide obtained from the organic extract of the tropical plant *Psychotropia longipes* that has been reported to have neurotensin inhibition properties [50]. Cpt A was able to inhibit neurotensin binding to its receptor to HT-29 cell membranes with an $IC_{50} \approx 3 \mu\text{M}$ and increase intracellular Ca^{2+} levels in a concentration-dependent manner, which could not be blocked by neurotensin antagonists [50]. Cpt A, however, showed a similar activity in two unrelated cell lines that did not express neurotensin receptors indicating that the mechanism of action is unlikely to be mediated through an interaction with the neurotensin receptor [50].

Chemical synthesis of cyclotides

Cyclotides are small peptides, approximately 30 amino acids long, and therefore can be readily synthesized by chemical methods using solid-phase peptide synthesis [51]. Chemical synthesis using a solid-phase approach has been utilized to generate native cyclotide structures as well as grafted analogues [52-56]. This method uses an intramolecular native chemical ligation [57], in which the peptide sequence contains an N-terminal cysteine and an α -thioester group at the C-terminus [58-60]. Both tert-butyloxycarbonyl (Boc)- and 9-fluorenyloxycarbonyl (Fmoc)-based chemistries have been used to incorporate C-terminal thioesters during chain assembly (Boc) [61-63] or using a safety-catch based linkers (Fmoc) [60, 64-67]. Once the peptide is cleaved from the resin, both cyclization and folding are carried out in a single pot reaction.

Recombinant expression of cyclotides

Cyclotides have also been produced recombinantly in bacteria through intramolecular native chemical ligation (see above) by using a modified protein splicing unit or intein (Fig. 3) (see reference [68] for a recent review). This method can generate folded cyclotides either *in vivo* or *in vitro* using standard bacterial expression systems [26, 69, 70]. Inteins are internal self-processing domains that undergo post-translational processing to splice together flanking external domains (exteins) [71]. The approach uses a modified intein fused to the C-terminus of the cyclotide sequence to allow the formation of an α -thioester at the C-terminus of recombinant polypeptides. To obtain the required N-terminal cysteine for cyclization, the peptide can be expressed with an N-terminal leading peptide signal, which can be cleaved either *in vivo* or *in vitro* by proteolysis or auto-proteolysis [68]. The simplest way to accomplish this is to introduce a Cys downstream of the initiating Met residue. Once the translation step is completed, the endogenous methionyl aminopeptidases (MAP) removes the Met residue, thereby generating *in vivo* an N-terminal Cys residue [72-76]. The N-

terminal Cys can then capture the reactive thioester in an intramolecular fashion to form a backbone-cyclized polypeptide (Fig. 3). Additional methods to generate an N-terminal cysteine have used exogenous proteases to cleave the leading signal after purification or *in vivo* by co-expressing the protease [77]. For example, the protease Factor Xa has been used to remove an N-terminal recognition sequence prior to a cysteine residue [59, 78]. Other proteases that have been used for this task include ubiquitin C-terminal hydrolase [79, 80], tobacco etch virus (TEV) protease [77], enterokinase [81] and thrombin [82]. The N-terminal pelB leader sequence has been used recently to direct newly synthesized fusion proteins to the *E. coli* periplasmic space where the corresponding endogenous leader peptidases [83, 84] can generate the desired N-terminal cysteine-containing protein fragment [85]. Besides proteases, protein splicing has also been used to produce recombinant N-terminal Cys-containing polypeptides. Some inteins can be modified in such a way that cleavage at the C-terminal splice junction can be accomplished in a pH- and temperature-dependent fashion [86-88].

Intein-mediated backbone cyclization of polypeptides has also been recently used for the biosynthesis of the Bowman-Birk inhibitor SFTI-1 [89]. The biosynthesis of other cyclic peptides such as backbone-cyclized α -defensins and naturally occurring θ -defensins is currently underway in our laboratory.

Another approach to generate cyclic peptides *in vivo* is by protein trans-splicing. This approach utilizes a self-processing intein that is split into two fragments, an N-intein and a C-intein. This method has not been applied yet for the biosynthesis of cyclotides, but has been used to produce other natural cyclic peptides and genetically-encoded libraries of small cyclic peptides [90, 91]. It should be noted, however, that these systems require the presence of specific amino acid residues at both intein-extein junctions for efficient protein splicing to occur [90, 92, 93].

Designing cyclotides with novel biological activities

The unique properties associated with the cyclotide scaffold make them extremely valuable tools in drug discovery. There are several studies that have used the cyclotide molecular scaffold to graft peptide sequences and to generate libraries for the purpose of engineering cyclotides with novel biological functions (Fig. 4).

The plasticity of the cyclotide framework was first demonstrated by substituting hydrophobic residues in loop 5 of kalata B1 with polar and charged residues [24]. The mutated cyclotides retained the native fold of kalata B1, but were no longer hemolytic [24]. This showed that cyclotides were amendable to sequence changes, and interestingly, can be modified to change their biological functions.

The potential of grafted cyclotides was first demonstrated in a study aimed to develop novel anticancer peptide-based therapeutics [94]. In this work a peptide antagonist of angiogenesis was grafted into various loops of the kalata B1 scaffold [94]. The grafted cyclotide containing the vascular endothelial growth factor A (VEGF-A) antagonist sequence in loop 3 was found to adopt a CCK native fold and was biologically active at low micromolar concentrations. Additionally, the grafted cyclotide showed increased resistance to degradation in human serum. This study demonstrates the possibility of using the cyclotide scaffold to stabilize bioactive peptide epitopes, which may normally get degraded.

The utility of the cyclotide scaffold in drug design has also been recently shown by engineering non-native activities into the cyclotide MCoTI-II. MCoTI-II is a naturally occurring trypsin inhibitor ($K_i \approx 20$ pM, which is the dissociation constant for inhibitor binding) found in the seeds of *Momordica cochinchinensis*, a tropical plant from the squash

family. Mutation of the P1 residue in the active loop (loop 1, see Fig. 1) of the cyclotide produced several MCoTI-II analogs with different specificities towards different proteases [95]. Interestingly, several analogs showed activity against the foot-and-mouth-disease virus (FMDV) 3C protease, a Cys protease key for viral replication, in the low micromolar range [95]. This is the first reported peptide-based inhibitor for this protease and although the potency was relatively low, this study demonstrates the potential of using MCoTI-based cyclotides for designing novel protease inhibitors [95].

In a more recent study, the same authors also generated inhibitors of the serine proteases β -tryptase and human leukocyte elastase (HLE) using the backbone of MCoTI-II [96]. β -Tryptase is implicated in allergic and inflammatory disorders, and HLE has been associated with respiratory and pulmonary disorders. Replacing the P1 residue in loop 1 produced several MCoTI-II mutants (K6A and K6V) with activity against HLE with K_i values of 20-30 nM [96] and K_i values against trypsin above 1 μ M. Removal of the SDGG peptide segment in loop 6 yielded a β -tryptase inhibitor with a $K_i \approx 10$ nM without significantly altering the three-dimensional structure as determined by NMR [96]. The authors hypothesized that deletion of the aspartic acid residue in MCoTI-II should improve activity by removal of repulsive electrostatic interactions with β -tryptase thus improving the inhibitory constant against trypsin 160-fold when compared to the wild-type MCoTI-II.

In addition to displaying biological activities, MCoTI-based peptides have also been shown to cross cell membranes in macrophage and breast cancer cell lines through macropinocytosis [21]. We have also found that grafting of a helix region from the molluscum contagiosum virus (MCV) FLICE-inhibitory protein (FLIP) into loop 6 of MCoTI-I yielded a folded cyclotide able to cross cell membranes and trigger apoptosis of virally infected cells (unpublished results), thus indicating that loop 6 may be used for grafting purposes without affecting cellular-uptake.

MCoTI cyclotides share a high sequence homology with related cystine-knot trypsin inhibitors found in squash, and could be considered cyclized homologs of these protease inhibitors. Squash cystine-knot trypsin inhibitors have also successfully been used to graft biological activities. For example, the RGD sequence was grafted into loop 1 of EETI-II yielding an EETI-II analog with platelet inhibitory activity [97]. The engineered proteins were much more potent in inhibiting platelet aggregation than the linear grafted peptides, thus highlighting the importance of grafting an epitope into a stable peptide-scaffold. These highly stable peptides could have clinical use for the treatment of patients with acute coronary syndrome, for example.

Additionally, the cyclotide scaffold can be engineered or evolved through molecular evolution techniques to selectively extracellular receptors such as binding G protein-coupled receptors to promote or block cell signaling [98]. These data demonstrate the versatility of the cyclotide scaffold and highlight the extraordinary pharmacological properties of MCoTI-cyclotides and related linear knottins, thus confirming the potential of Cys-knotted polypeptide scaffolds in peptide-based drug discovery [27].

Screening of cyclotide-based libraries

The ability to create cyclic polypeptides *in vivo* opens up the intriguing possibility of generating large libraries of cyclic polypeptides. Thus, libraries of genetically-encoded cyclic polypeptides containing billions of members can be readily generated using standard molecular recombinant tools. This tremendous molecular diversity allows one to perform selection strategies mimicking the evolutionary processes found in nature.

There are several examples where *in vivo* generated libraries of cyclic peptides have been used for rapid selection of biologically active peptides. For example protein trans-splicing has been used by several groups for the generation of libraries of small cyclic peptides (≤ 8 residues) in bacterial and mammalian cells [90, 92, 93, 99-101]. It should be pointed out, however, that the use of protein trans-splicing requires particular amino acids at the intein-extein junctions for efficient trans-splicing [93]. These requirements usually depend on the type of split-intein used, and seriously limit the diversity of the libraries that can be generated when using small cyclic peptide templates. Our group has recently demonstrated the expression in *E. coli* of libraries based on the cyclic peptide SFTI-1 (a backbone cyclized Bowman-Birk trypsin inhibitor) using an intein-mediated backbone cyclization approach (see above), which allows the biosynthesis of backbone-cyclized polypeptides without any sequence requirement limitation [89].

The use of small cyclic peptides, however, which technically can be considered a single closed loop (or 2 loops in the case of SFTI-1), could also limit the potency of the peptides selected, especially when targeting protein-protein interactions involving large binding surfaces. In these cases the use of peptide templates such as cyclotides with multiple variable loops could facilitate the selection of peptides with higher affinities.

The potential for generating cyclotide libraries was first explored by our group using the kalata B1 scaffold [69]. In this work wild-type and several mutants of kalata B1 were biosynthesized using an intramolecular native chemical ligation facilitated by a modified protein splicing unit. In this work, six different linear versions of kalata B1 were generated and expressed in *E. coli* as fusions to a modified version of the yeast vacuolar membrane ATPase (VMA) intein. Results demonstrated *in vitro* folding and cyclization of kalata B1 to varying degrees depending on which of the six native cysteine residues was at the N-terminus after cleavage of the initiation methionine by endogenous MAP. Cleavage and efficient cyclization of the different linear precursors did not occur equally, suggesting the amino acid residues near the intein as well as the predisposition to adopt a native fold of the corresponding linear precursor may determine the efficiency of the cleavage/cyclization step [69]. This information was used to express a small library based on the kalata B1 scaffold. This library was cyclized *in vitro* by incubation with a redox buffer containing reduced glutathione (GSH) as a thiol co-factor, thus mimicking the intracellular conditions, where GSH is the most abundant thiol co-factor. The use of GSH allows the cyclization and folding to happen in one step [69, 89]. Analysis of the cyclization/folding reaction by HPLC and mass spectrometry revealed that all the members of the kalata B1 based library were expressed and processed with similar yields to give the corresponding natively folded cyclotides [69].

More recently, we have also reported the biosynthesis of a genetically encoded library of MCoTI-I based cyclotides in *E. coli* cells [26]. The cyclization/folding of the library was performed either *in vitro*, by incubation with a redox buffer containing glutathione, or by *in vivo* self-processing of the corresponding precursor proteins. The bacterial gyrase A intein from *Mycobacterium xenopus* was used in this work [26]. This intein typically express at higher yields than the yeast VMA intein in *E. coli* expression systems [26]. The peptide libraries were purified and screened for activity using trypsin-immobilized sepharose beads, and then analyzed by HPLC and mass spectrometry. Out of 27 mutations studied, only two mutations, G27P and I22G, negatively affected the folding of the resulting cyclotides. All of the remaining cyclotides were able to fold with similar yields. The K6A mutant, as expected, was not able to bind trypsin. This residue is key for binding to the specificity pocket of trypsin, and can only be replaced by positively charged residues [102]. This mutant was found by NMR to adopt a native cyclotide structure, confirming that the lack of biological activity was due to the mutation and not to the ability to adopt a native fold. It is interesting

to note that by modifying the nature of this residue, the specificity of the corresponding MCoTI-cyclotide can be changed to target other proteases [95, 96]. The rest of the MCoTI-based library members were able to bind trypsin, suggesting they were able to adopt a native cyclotide fold and retained biological activity. The affinity of each peptide was analyzed using a competitive trypsin-binding assay. The mutants had a wide range of affinity, some being greater than wild type MCoTI-I. The peptides with less affinity were mostly found in loop 1 and the C-terminal region of loop 6, both well conserved among other squash trypsin inhibitors. Overall, these data describe the structural requirements for correct formation of MCoTI-I and the residues that are key to modulating trypsin binding. To our knowledge, this is the first time that the biosynthesis of a genetically-encoded library of MCoTI-based cyclotides containing a complete suite of amino acid mutants is reported.

The chemical synthesis of a complete suite of Ala mutants for kalata B1 has also been recently reported [25]. In this work all the mutants were fully characterized structurally and functionally. The results indicated that only two of the mutations explored (W23A and P24A, both located in loop 5, see Fig. 1) prevented folding [25]. The mutagenesis results obtained in our work with the cyclotide MCoTI-I show similar results highlighting the extreme robustness of the cyclotide scaffold to mutations. These studies show that cyclotides may provide an ideal scaffold for the biosynthesis of large combinatorial libraries inside living bacterial cells. These genetically-encoded libraries can then be screened in-cell for biological activity using high-throughput flow cytometry techniques for the rapid selection of novel biologically active cyclotides [69, 103, 104].

Summary and Concluding Remarks

In summary, cyclotides are a novel family of structurally related globular microproteins with a unique head-to-tail cyclized backbone, which is stabilized by three disulfide bonds [27, 105]. The number and positions of cysteine residues are conserved throughout the family, forming what is called cyclic cystine-knot (CCK) motif [35] that acts as a highly stable and versatile scaffold on which 5 hyper-variable loops are arranged (Fig 1). This CCK framework gives the cyclotides exceptional resistance to thermal and chemical denaturation, and enzymatic degradation. This is particularly important for the development of peptide-based therapeutics with oral bioavailability. In fact, the use of cyclotide-containing plants in indigenous medicine first highlighted the fact that the peptides are resistant to boiling and are apparently orally bioavailable [14, 32, 33]. Some cyclotides have also been shown to cross the cell membrane [21] thus allowing to target intracellular protein interactions, such as that mediated by viral FLIPs to prevent cell-death in virally infected cells (unpublished results). Cyclotides are also medium-sized polypeptides and therefore can be readily synthesized by standard solid-phase peptide-synthesis using either Boc-[54] or Fmoc-based [55] methodologies thus allowing the introduction of non-natural amino acids or other chemical modifications for lead optimization. They can also be encoded within standard cloning vectors and readily expressed in bacteria or animal cells [26, 70], thus making them ideal substrates for molecular evolution strategies to enable generation and selection of compounds with optimal binding and inhibitory characteristics using high throughput cell-based assays [106]. All of these characteristics make cyclotides appear as very promising leads or frameworks for development of peptide-based therapeutics and diagnostics [27, 68, 105, 107].

Acknowledgments

This work was supported by funding from the School of Pharmacy at the University of Southern California and by National Institute of Health award GM090323 to J.A.C.

Abbreviations

AEP	asparaginyl endoproteinase
Boc	tert-butyloxycarbonyl
EETI-II	<i>Ecballium elaterium</i> trypsin inhibitor II
FMDV	foot-and-mouth-disease virus
FLIP	FLICE-inhibitory protein
Fmoc	9-fluorenyloxycarbonyl
GSH	reduced glutathione
HIV	human immunodeficiency virus
HLE	human leukocyte elastase
HPLC	high performance liquid chromatography
MAP	methionyl aminopeptidase
MCoTI	<i>Momordica cochinchinensis</i> trypsin inhibitor
MCV	molluscum contagiosum virus
NMR	nuclear magnetic resonance
NTR	N-terminal repeat
SFTI	sunflower trypsin inhibitor
SPPS	solid-phase peptide synthesis
VMA	vacuolar membrane ATPase

References

- [1]. Mogi T, Kita K. Gramicidin S and polymyxins: the revival of cationic cyclic peptide antibiotics. *Cell Mol Life Sci.* 2009; 66:3821–3826. [PubMed: 19701717]
- [2]. Starzl TE, Klintmalm GB, Porter KA, Iwatsuki S, Schroter GP. Liver transplantation with use of cyclosporin a and prednisone. *N Engl J Med.* 1981; 305:266–269. [PubMed: 7017414]
- [3]. Enoch DA, Bygott JM, Daly ML, Karas JA. Daptomycin. *J Infect.* 2007; 55:205–213. [PubMed: 17629567]
- [4]. Tang YQ, Yuan J, Osapay G, Osapay K, Tran D, Miller CJ, Ouellette AJ, Selsted ME. A cyclic antimicrobial peptide produced in primate leukocytes by the ligation of two truncated alpha-defensins. *Science.* 1999; 286:498–502. [PubMed: 10521339]
- [5]. Garcia AE, Osapay G, Tran PA, Yuan J, Selsted ME. Isolation, synthesis, and antimicrobial activities of naturally occurring theta-defensin isoforms from baboon leukocytes. *Infect Immun.* 2008; 76:5883–5891. [PubMed: 18852242]
- [6]. Gallo SA, Wang W, Rawat SS, Jung G, Waring AJ, Cole AM, Lu H, Yan X, Daly NL, Craik DJ, Jiang S, Lehrer RI, Blumenthal R. Theta-defensins prevent HIV-1 Env-mediated fusion by binding gp41 and blocking 6-helix bundle formation. *J Biol Chem.* 2006; 281:18787–18792. [PubMed: 16648135]
- [7]. Wang W, Mulakala C, Ward SC, Jung G, Luong H, Pham D, Waring AJ, Kaznessis Y, Lu W, Bradley KA, Lehrer RI. Retrocyclins kill bacilli and germinating spores of *Bacillus anthracis* and inactivate anthrax lethal toxin. *J Biol Chem.* 2006; 281:32755–32764. [PubMed: 16790431]
- [8]. Seidel A, Ye Y, de Armas LR, Soto M, Yarosh W, Marcisin RA, Tran D, Selsted ME, Camerini D. Cyclic and acyclic defensins inhibit human immunodeficiency virus type-1 replication by different mechanisms. *PLoS One.* 2010; 5:e9737. [PubMed: 20305815]

- [9]. Craik DJ. Circling the enemy: cyclic proteins in plant defence. *Trends Plant Sci.* 2009; 14:328–335. [PubMed: 19423383]
- [10]. Korsinczky ML, Schirra HJ, Craik DJ. Sunflower trypsin inhibitor-1. *Curr Protein Pept Sci.* 2004; 5:351–364. [PubMed: 15544530]
- [11]. Wang CK, Kaas Q, Chiche L, Craik DJ. CyBase: a database of cyclic protein sequences and structures, with applications in protein discovery and engineering. *Nucleic Acids Res.* 2008; 36:D206–210. [PubMed: 17986451]
- [12]. Jennings C, West J, Waine C, Craik D, Anderson M. Biosynthesis and insecticidal properties of plant cyclotides: the cyclic knotted proteins from Oldenlandia affinis. *Proc Natl Acad Sci U S A.* 2001; 98:10614–10619. [PubMed: 11535828]
- [13]. Jennings CV, Rosengren KJ, Daly NL, Plan M, Stevens J, Scanlon MJ, Waine C, Norman DG, Anderson MA, Craik DJ. Isolation, solution structure, and insecticidal activity of kalata B2, a circular protein with a twist: do Möbius strips exist in nature? *Biochemistry.* 2005; 44:851–860. [PubMed: 15654741]
- [14]. Gran L. Oxytocic principles of Oldenlandia affinis. *Lloydia.* 1973; 36:174–178. [PubMed: 4744554]
- [15]. Gustafson KR, Sowder RC, Henderson LE, Parsons IC, Kashman Y, Cardellina JH, McMahon JB, Buckheit RW, Pannell LK, Boyd MR. Circulin-A and Circulin-B: Novel HIV-inhibitory macrocyclic peptides from the tropical tree Chassalia parvifolia. *J Am Chem Soc.* 1994; 116:9337–9338.
- [16]. Tam JP, Lu YA, Yang JL, Chiu KW. An unusual structural motif of antimicrobial peptides containing end-to-end macrocycle and cystine-knot disulfides. *Proc Natl Acad Sci U S A.* 1999; 96:8913–8918. [PubMed: 10430870]
- [17]. Gran L, Sletten K, Skjeldal L. Cyclic peptides from Oldenlandia affinis DC. Molecular and biological properties. *Chem Biodivers.* 2008; 5:2014–2022. [PubMed: 18972522]
- [18]. Lindholm P, Goransson U, Johansson S, Claeson P, Gullbo J, Larsson R, Bohlin L, Backlund A. Cyclotides: a novel type of cytotoxic agents. *Mol Cancer Ther.* 2002; 1:365–369. [PubMed: 12477048]
- [19]. Colgrave ML, Kotze AC, Ireland DC, Wang CK, Craik DJ. The anthelmintic activity of the cyclotides: natural variants with enhanced activity. *Chembiochem.* 2008; 9:1939–1945. [PubMed: 18618891]
- [20]. Colgrave ML, Kotze AC, Huang YH, O’Grady J, Simonsen SM, Craik DJ. Cyclotides: natural, circular plant peptides that possess significant activity against gastrointestinal nematode parasites of sheep. *Biochemistry.* 2008; 47:5581–5589. [PubMed: 18426225]
- [21]. Greenwood KP, Daly NL, Brown DL, Stow JL, Craik DJ. The cyclic cystine knot miniprotein MCoTI-II is internalized into cells by macropinocytosis. *Int J Biochem Cell Biol.* 2007; 39:2252–2264. [PubMed: 17693122]
- [22]. Colgrave ML, Craik DJ. Thermal, chemical, and enzymatic stability of the cyclotide kalata B1: the importance of the cyclic cystine knot. *Biochemistry.* 2004; 43:5965–5975. [PubMed: 15147180]
- [23]. Craik DJ, Cemazar M, Wang CK, Daly NL. The cyclotide family of circular miniproteins: nature’s combinatorial peptide template. *Biopolymers.* 2006; 84:250–266. [PubMed: 16440288]
- [24]. Clark RJ, Daly NL, Craik DJ. Structural plasticity of the cyclic-cystine-knot framework: implications for biological activity and drug design. *Biochem J.* 2006; 394:85–93. [PubMed: 16300479]
- [25]. Simonsen SM, Sando L, Rosengren KJ, Wang CK, Colgrave ML, Daly NL, Craik DJ. Alanine scanning mutagenesis of the prototypic cyclotide reveals a cluster of residues essential for bioactivity. *J Biol Chem.* 2008; 283:9805–9813. [PubMed: 18258598]
- [26]. Austin J, Wang W, Puttamadappa S, Shekhtman A, Camarero JA. Biosynthesis and biological screening of a genetically encoded library based on the cyclotide MCoTI-I. *Chembiochem.* 2009; 10:2663–2670. [PubMed: 19780078]
- [27]. Jagadish K, Camarero JA. Cyclotides, a promising molecular scaffold for peptide-based therapeutics. *Biopolymers.* 2010

[28]. Gillon AD, Saska I, Jennings CV, Guarino RF, Craik DJ, Anderson MA. Biosynthesis of circular proteins in plants. *Plant J.* 2008; 53:505–515. [PubMed: 18086282]

[29]. Saska I, Gillon AD, Hatsugai N, Dietzgen RG, Hara-Nishimura I, Anderson MA, Craik DJ. An asparaginyl endopeptidase mediates in vivo protein backbone cyclization. *J Biol Chem.* 2007; 282:29721–29728. [PubMed: 17698845]

[30]. Sheldon PS, Keen JN, Bowles DJ. Post-translational peptide bond formation during concanavalin A processing in vitro. *Biochem J.* 1996; 320(Pt 3):865–870. [PubMed: 9003373]

[31]. Plan MR, Saska I, Cagauan AG, Craik DJ. Backbone cyclised peptides from plants show molluscicidal activity against the rice pest *Pomacea canaliculata* (golden apple snail). *J Agric Food Chem.* 2008; 56:5237–5241. [PubMed: 18557620]

[32]. Gran L. On the effect of a polypeptide isolated from “Kalata-Kalata” (*Oldenlandia affinis* DC) on the oestrogen dominated uterus. *Acta Pharmacol Toxicol (Copenh).* 1973; 33:400–408. [PubMed: 4801084]

[33]. Gran L, Sandberg F, Sletten K. *Oldenlandia affinis* (R&S) DC. A plant containing uteroactive peptides used in African traditional medicine. *J Ethnopharmacol.* 2000; 70:197–203. [PubMed: 10837983]

[34]. Saether O, Craik DJ, Campbell ID, Sletten K, Juul J, Norman DG. Elucidation of the primary and three-dimensional structure of the uterotonic polypeptide kalata B1. *Biochemistry.* 1995; 34:4147–4158. [PubMed: 7703226]

[35]. Craik DJ, Daly NL, Bond T, Waine C. Plant cyclotides: A unique family of cyclic and knotted proteins that defines the cyclic cystine knot structural motif. *J Mol Biol.* 1999; 294:1327–1336. [PubMed: 10600388]

[36]. Goransson U, Luijendijk T, Johansson S, Bohlin L, Claeson P. Seven novel macrocyclic polypeptides from *Viola arvensis*. *J Nat Prod.* 1999; 62:283–286. [PubMed: 10075760]

[37]. Henriques ST, Craik DJ. Cyclotides as templates in drug design. *Drug Discov Today.* 15:57–64. [PubMed: 19878736]

[38]. Daly NL, Gustafson KR, Craik DJ. The role of the cyclic peptide backbone in the anti-HIV activity of the cyclotide kalata B1. *FEBS Lett.* 2004; 574:69–72. [PubMed: 15358541]

[39]. Chen B, Colgrave ML, Daly NL, Rosengren KJ, Gustafson KR, Craik DJ. Isolation and characterization of novel cyclotides from *Viola hederaceae*: solution structure and anti-HIV activity of vhl-1, a leaf-specific expressed cyclotide. *J Biol Chem.* 2005; 280:22395–22405. [PubMed: 15824119]

[40]. Wang CK, Colgrave ML, Gustafson KR, Ireland DC, Goransson U, Craik DJ. Anti-HIV cyclotides from the Chinese medicinal herb *Viola yedoensis*. *J Nat Prod.* 2008; 71:47–52. [PubMed: 18081258]

[41]. Kamimori H, Hall K, Craik DJ, Aguilar MI. Studies on the membrane interactions of the cyclotides kalata B1 and kalata B6 on model membrane systems by surface plasmon resonance. *Anal Biochem.* 2005; 337:149–153. [PubMed: 15649388]

[42]. Shenkarev ZO, Nadezhdin KD, Sobol VA, Sobol AG, Skjeldal L, Arseniev AS. Conformation and mode of membrane interaction in cyclotides. Spatial structure of kalata B1 bound to a dodecylphosphocholine micelle. *FEBS J.* 2006; 273:2658–2672. [PubMed: 16817894]

[43]. Shenkarev ZO, Nadezhdin KD, Lyukmanova EN, Sobol VA, Skjeldal L, Arseniev AS. Divalent cation coordination and mode of membrane interaction in cyclotides: NMR spatial structure of ternary complex Kalata B7/Mn²⁺/DPC micelle. *J Inorg Biochem.* 2008; 102:1246–1256. [PubMed: 18295894]

[44]. Wang CK, Hu SH, Martin JL, Sjogren T, Hajdu J, Bohlin L, Claeson P, Goransson U, Rosengren KJ, Tang J, Tan NH, Craik DJ. Combined X-ray and NMR analysis of the stability of the cyclotide cystine knot fold that underpins its insecticidal activity and potential use as a drug scaffold. *J Biol Chem.* 2009; 284:10672–10683. [PubMed: 19211551]

[45]. Svengard E, Goransson U, Hocaoglu Z, Gullbo J, Larsson R, Claeson P, Bohlin L. Cytotoxic cyclotides from *Viola tricolor*. *J Nat Prod.* 2004; 67:144–147. [PubMed: 14987049]

[46]. Herrmann A, Burman R, Mylne JS, Karlsson G, Gullbo J, Craik DJ, Clark RJ, Goransson U. The alpine violet, *Viola biflora*, is a rich source of cyclotides with potent cytotoxicity. *Phytochemistry.* 2008; 69:939–952. [PubMed: 18191970]

[47]. Herrmann A, Svangard E, Claeson P, Gullbo J, Bohlin L, Goransson U. Key role of glutamic acid for the cytotoxic activity of the cyclotide cycloviolacin O2. *Cell Mol Life Sci.* 2006; 63:235–245. [PubMed: 16389447]

[48]. Chen B, Colgrave ML, Wang C, Craik DJ. Cycloviolacin H4, a hydrophobic cyclotide from *Viola hederacea*. *J Nat Prod.* 2006; 69:23–28. [PubMed: 16441062]

[49]. Huang YH, Colgrave ML, Clark RJ, Kotze AC, Craik DJ. Lysine-scanning mutagenesis reveals an amendable face of the cyclotide kalata B1 for the optimization of nematocidal activity. *J Biol Chem.* 2010; 285:10797–10805. [PubMed: 20103593]

[50]. Witherup KM, Bogusky MJ, Anderson PS, Ramjit H, Ransom RW, Wood T, Sardana M. Cyclopsychotride A, a biologically active, 31-residue cyclic peptide isolated from *Psychotria longipes*. *J Nat Prod.* 1994; 57:1619–1625. [PubMed: 7714530]

[51]. Marglin A, Merrifield RB. Chemical synthesis of peptides and proteins. *Annu Rev Biochem.* 1970; 39:841–866. [PubMed: 4920831]

[52]. Tam JP, Lu Y-A. Synthesis of Large Cyclic Cystine-Knot Peptide by Orthogonal Coupling Strategy Using Unprotected Peptide Precursor. *Tetrahedron Lett.* 1997; 38:5599–5602.

[53]. Tam JP, Lu YA. A biomimetic strategy in the synthesis and fragmentation of cyclic protein. *Protein Sci.* 1998; 7:1583–1592. [PubMed: 9684891]

[54]. Daly NL, Love S, Alewood PF, Craik DJ. Chemical synthesis and folding pathways of large cyclic polypeptides: studies of the cystine knot polypeptide kalata B1. *Biochemistry.* 1999; 38:10606–10614. [PubMed: 10441158]

[55]. Thongyoo P, Tate EW, Leatherbarrow RJ. Total synthesis of the macrocyclic cysteine knot microprotein MCoTI-II. *Chem Commun (Camb).* 2006:2848–2850. [PubMed: 17007393]

[56]. Aboye, T. Leta; Clark, RJ.; Craik, DJ.; Goransson, U. Ultra-stable peptide scaffolds for protein engineering-synthesis and folding of the circular cystine knotted cyclotide cycloviolacin O2. *Chembiochem.* 2008; 9:103–113. [PubMed: 18058973]

[57]. Dawson PE, Muir TW, Clark-Lewis I, Kent SBH. Synthesis of Proteins by Native Chemical Ligation. *Science.* 1994; 266:776–779. [PubMed: 7973629]

[58]. Camarero JA, Muir TW. Chemoselective backbone cyclization of unprotected peptides. *J. Chem. Soc., Chem. Comm.* 1997; 1997:1369–1370.

[59]. Camarero JA, Muir TW. Biosynthesis of a Head-to-Tail Cyclized Protein with Improved Biological Activity. *J. Am. Chem. Soc.* 1999; 121:5597–5598.

[60]. Camarero JA, Mitchell AR. Synthesis of proteins by native chemical ligation using Fmoc-based chemistry. *Protein Pept Lett.* 2005; 12:723–728. [PubMed: 16305540]

[61]. Camarero JA, Cotton GJ, Adeva A, Muir TW. Chemical ligation of unprotected peptides directly from a solid support. *J Pept Res.* 1998; 51:303–316. [PubMed: 9560006]

[62]. Beligere GS, Dawson PE. Conformationally assisted ligation using C-terminal thioester peptides. *J. Am. Chem. Soc.* 1999; 121:6332–6333.

[63]. Camarero JA, Adeva A, Muir TW. 3-Thiopropionic acid as a highly versatile multidetachable thioester resin linker. *Lett. Pept. Sci.* 2000; 7:17–21.

[64]. Ingenito R, Bianchi E, Fattori D, Pessi A. Solid phase synthesis of peptide C-terminal thioesters by Fmoc/t-Bu chemistry Source. *J. Am. Chem. Soc.* 1999; 121:11369–11374.

[65]. Shin Y, Winans KA, Backes BJ, Kent SBH, Ellman JA, Bertozzi CR. Fmoc-Based Synthesis of Peptide-^aThioesters: Application to the Total Chemical Synthesis of a Glycoprotein by Native Chemical Ligation. *J. Am. Chem. Soc.* 1999; 121:11684–11689.

[66]. Camarero JA, Hackel BJ, de Yoreo JJ, Mitchell AR. Fmoc-based synthesis of peptide alpha-thioesters using an aryl hydrazine support. *J Org Chem.* 2004; 69:4145–4151. [PubMed: 15176841]

[67]. Blanco-Canosa JB, Dawson PE. An efficient Fmoc-SPPS approach for the generation of thioester peptide precursors for use in native chemical ligation. *Angew Chem Int Ed Engl.* 2008; 47:6851–6855. [PubMed: 18651678]

[68]. Sancheti H, Camarero JA. “Splicing up” drug discovery. Cell-based expression and screening of genetically-encoded libraries of backbone-cyclized polypeptides. *Adv Drug Deliv Rev.* 2009; 61:908–917. [PubMed: 19628015]

[69]. Kimura RH, Tran AT, Camarero JA. Biosynthesis of the cyclotide kalata B1 by using protein splicing. *Angewandte Chemie (International ed.)*. 2006; 45:973–976.

[70]. Camarero JA, Kimura RH, Woo YH, Shekhtman A, Cantor J. Biosynthesis of a fully functional cyclotide inside living bacterial cells. *Chembiochem*. 2007; 8:1363–1366. [PubMed: 17590879]

[71]. Perler FB, Adam E. Protein splicing and its applications. *Curr. Opin. Biotechnol.* 2000;377–383. [PubMed: 10975457]

[72]. Hirel PH, Schmitter MJ, Dessen P, Fayat G, Blanquet S. Extent of N-terminal methionine excision from *Escherichia coli* proteins is governed by the side-chain length of the penultimate amino acid. *Proc. Natl. Acad. Sci. U S A.* 1989; 86:8247–8251. [PubMed: 2682640]

[73]. Dwyer MA, Lu W, Dwyer JJ, Kossiakoff AA. Biosynthetic phage display: a novel protein engineering tool combining chemical and genetic diversity. *Chem. Biol.* 2000; 7:263–274. [PubMed: 10780926]

[74]. Iwai H, Pluckthum A. Circular b-lactamase: stability enhancement by cyclizing the backbone. *FEBS Lett.* 1999;166–172. [PubMed: 10518012]

[75]. Camarero JA, Fushman D, Cowburn D, Muir TW. Peptide chemical ligation inside living cells: in vivo generation of a circular protein domain. *Bioorg Med Chem.* 2001; 9:2479–2484. [PubMed: 11553489]

[76]. Cotton GJ, Ayers B, Xu R, Muir TW. Insertion of a Synthetic Peptide into a Recombinant Protein Framework; A Protein Biosensor. *J. Am. Chem. Soc.* 1999; 121:1100–1101.

[77]. Tolbert TJ, Wong C-H. New methods for proteomic research: preparation of proteins with N-terminal cysteines for labeling and conjugation. *Angew. Chem. Int. Ed.* 2002; 41:2171–2174.

[78]. Erlanson DA, Chytil M, Verdine GL. The leucine zipper domain controls the orientation of AP-1 in the NFAT/AP-1.DNA complex. *Chem Biol.* 1996; 3:981–991. [PubMed: 9000009]

[79]. Baker RT, Smith SA, Marano R, McKee J, Board PG. Protein expression using cotranslational fusion and cleavage of ubiquitin. Mutagenesis of the glutathione-binding site of human Pi class glutathione S-transferase. *J Biol Chem.* 1994; 269:25381–25386. [PubMed: 7929235]

[80]. Varshavsky A. Ubiquitin fusion technique and related methods. *Methods Enzymol.* 2005; 399:777–799. [PubMed: 16338395]

[81]. Dykes CW, Bookless AB, Coomber BA, Noble SA, Humber DC, Hobden AN. Expression of atrial natriuretic factor as a cleavable fusion protein with chloramphenicol acetyltransferase in *Escherichia coli*. *Eur J Biochem.* 1988; 174:411–416. [PubMed: 2968246]

[82]. Liu D, Xu R, Dutta K, Cowburn D. N-terminal cysteinyl proteins can be prepared using thrombin cleavage. *FEBS Lett.* 2008; 582:1163–1167. [PubMed: 18331839]

[83]. Dalbey RE, Lively MO, Bron S, van Dijl JM. The chemistry and enzymology of the type I signal peptidases. *Protein Sci.* 1997; 6:1129–1138. [PubMed: 9194173]

[84]. Paetzel M, Dalbey RE, Strynadka NC. Crystal structure of a bacterial signal peptidase apoenzyme: implications for signal peptide binding and the Ser-Lys dyad mechanism. *J Biol Chem.* 2002; 277:9512–9519. [PubMed: 11741964]

[85]. Hauser PS, Ryan RO. Expressed protein ligation using an N-terminal cysteine containing fragment generated in vivo from a pelB fusion protein. *Protein Expr Purif.* 2007; 54:227–233. [PubMed: 17493830]

[86]. Evans TC, Benner J, Xu M-Q. The in Vitro Ligation of Bacterially Expressed Proteins Using an Intein from *Metanobacterium thermoautotrophicum*. *J. Biol. Chem.* 1999; 274:3923–3926. [PubMed: 9933578]

[87]. Southworth MW, Amaya K, Evans TC, Xu MQ, Perler FB. Purification of proteins fused to either the amino or carboxy terminus of the *Mycobacterium xenopi* gyrase A intein. *Biotechniques.* 1999; 27116:110–114. 118–120. [PubMed: 10407673]

[88]. Mathys S, Evans TC, Chute IC, Wu H, Chong S, Benner J, Liu XQ, Xu MQ. Characterization of a self-splicing mini-intein and its conversion into autocatalytic N- and C-terminal cleavage elements: facile production of protein building blocks for protein ligation. *Gene.* 1999; 231:1–13. [PubMed: 10231563]

[89]. Austin J, Kimura RH, Woo YH, Camarero JA. In vivo biosynthesis of an Ala-scan library based on the cyclic peptide SFTI-1. *Amino Acids.* 2010; 38:1313–1322. [PubMed: 19685144]

[90]. Abel-Santos E, Scott CP, Benkovic SJ. Use of inteins for the in vivo production of stable cyclic peptide libraries in *E. coli*. *Methods Mol. Biol.* 2003; 205:281–294. [PubMed: 12491894]

[91]. Tavassoli A, Benkovic SJ. Split-intein mediated circular ligation used in the synthesis of cyclic peptide libraries in *E. coli*. *Nat Protoc.* 2007; 2:1126–1133. [PubMed: 17546003]

[92]. Scott CP, Abel-Santos E, Wall M, Wahnon D, Benkovic SJ. Production of cyclic peptides and proteins in vivo. *Proc. Natl. Acad. Sci. USA.* 1999; 96:13638–13643. [PubMed: 10570125]

[93]. Scott CP, Abel-Santos E, Jones AD, Benkovic SJ. Structural requirements for the biosynthesis of backbone cyclic peptide libraries. *Chem. Biol.* 2001; 8:801–815. [PubMed: 11514229]

[94]. Gunasekera S, Foley FM, Clark RJ, Sando L, Fabri LJ, Craik DJ, Daly NL. Engineering stabilized vascular endothelial growth factor-A antagonists: synthesis, structural characterization, and bioactivity of grafted analogues of cyclotides. *J Med Chem.* 2008; 51:7697–7704. [PubMed: 19053834]

[95]. Thongyoo P, Roque-Rosell N, Leatherbarrow RJ, Tate EW. Chemical and biomimetic total syntheses of natural and engineered MCoTI cyclotides. *Org Biomol Chem.* 2008; 6:1462–1470. [PubMed: 18385853]

[96]. Thongyoo P, Bonomelli C, Leatherbarrow RJ, Tate EW. Potent inhibitors of beta-tryptase and human leukocyte elastase based on the MCoTI-II scaffold. *J Med Chem.* 2009; 52:6197–6200. [PubMed: 19772295]

[97]. Reiss S, Sieber M, Oberle V, Wentzel A, Spangenberg P, Claus R, Kolmar H, Losche W. Inhibition of platelet aggregation by grafting RGD and KGD sequences on the structural scaffold of small disulfide-rich proteins. *Platelets.* 2006; 17:153–157. [PubMed: 16702041]

[98]. Gruber CW, Muttenthaler M, Freissmuth M. Ligand-Based Peptide Design and Combinatorial Peptide Libraries to Target G Protein-Coupled Receptors. *Curr Pharm Des.* 2010

[99]. Kinsella TM, Ohashi CT, Harder AG, Yam GC, Li W, Peelle B, Pali ES, Bennett MK, Molineaux SM, Anderson DA, Masuda ES, Payan DG. Retrovirally delivered random cyclic Peptide libraries yield inhibitors of interleukin-4 signaling in human B cells. *J Biol Chem.* 2002; 277:37512–37518. [PubMed: 12167667]

[100]. Cheng L, Naumann TA, Horswill AR, Hong SJ, Venters BJ, Tomsho JW, Benkovic SJ, Keiler KC. Discovery of antibacterial cyclic peptides that inhibit the ClpXP protease. *Protein Sci.* 2007; 16:1535–1542. [PubMed: 17600141]

[101]. Tavassoli A, Lu Q, Gam J, Pan H, Benkovic SJ, Cohen SN. Inhibition of HIV budding by a genetically selected cyclic peptide targeting the Gag-TSG101 interaction. *ACS Chem Biol.* 2008; 3:757–764. [PubMed: 19053244]

[102]. Hernandez JF, Gagnon J, Chiche L, Nguyen TM, Andrieu JP, Heitz A, Hong T, Trinh, Pham TT, Le Nguyen D. Squash trypsin inhibitors from *Momordica cochinchinensis* exhibit an atypical macrocyclic structure. *Biochemistry.* 2000; 39:5722–5730. [PubMed: 10801322]

[103]. Giepmans BN, Adams SR, Ellisman MH, Tsien RY. The fluorescent toolbox for assessing protein location and function. *Science.* 2006; 312:217–224. [PubMed: 16614209]

[104]. You X, Nguyen AW, Jabaiah A, Sheff MA, Thorn KS, Daugherty PS. Intracellular protein interaction mapping with FRET hybrids. *Proc Natl Acad Sci U S A.* 2006; 103:18458–18463. [PubMed: 17130455]

[105]. Daly NL, Rosengren KJ, Craik DJ. Discovery, structure and biological activities of cyclotides. *Adv Drug Deliv Rev.* 2009; 61:918–930. [PubMed: 19470399]

[106]. Kimura RH, Steenblock ER, Camarero JA. Development of a cell-based fluorescence resonance energy transfer reporter for *Bacillus anthracis* lethal factor protease. *Anal Biochem.* 2007; 369:60–70. [PubMed: 17586456]

[107]. Kolmar H. Biological diversity and therapeutic potential of natural and engineered cystine knot miniproteins. *Curr Opin Pharmacol.* 2009; 9:608–614. [PubMed: 19523876]

[108]. Rosengren KJ, Daly NL, Plan MR, Waine C, Craik DJ. Twists, knots, and rings in proteins. Structural definition of the cyclotide framework. *J Biol Chem.* 2003; 278:8606–8616. [PubMed: 12482868]

[109]. Felizmenio-Quimio ME, Daly NL, Craik DJ. Circular proteins in plants: solution structure of a novel macrocyclic trypsin inhibitor from *Momordica cochinchinensis*. *J Biol Chem.* 2001; 276:22875–22882. [PubMed: 11292835]

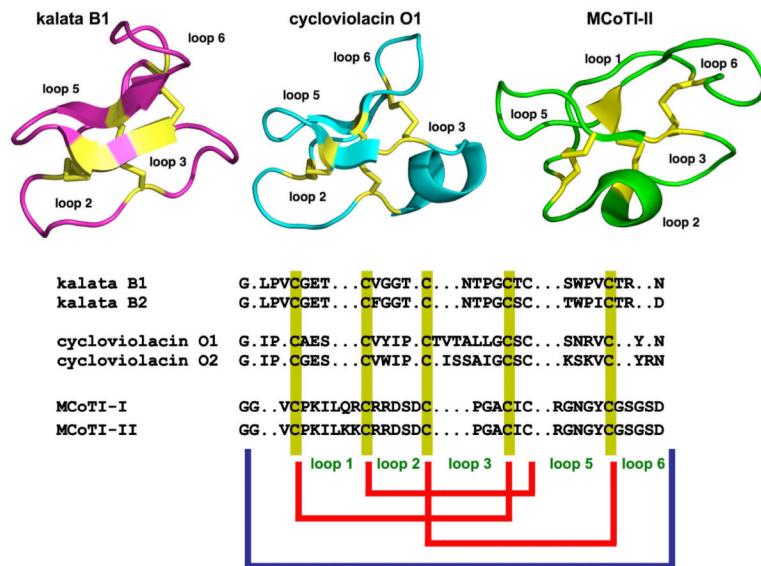
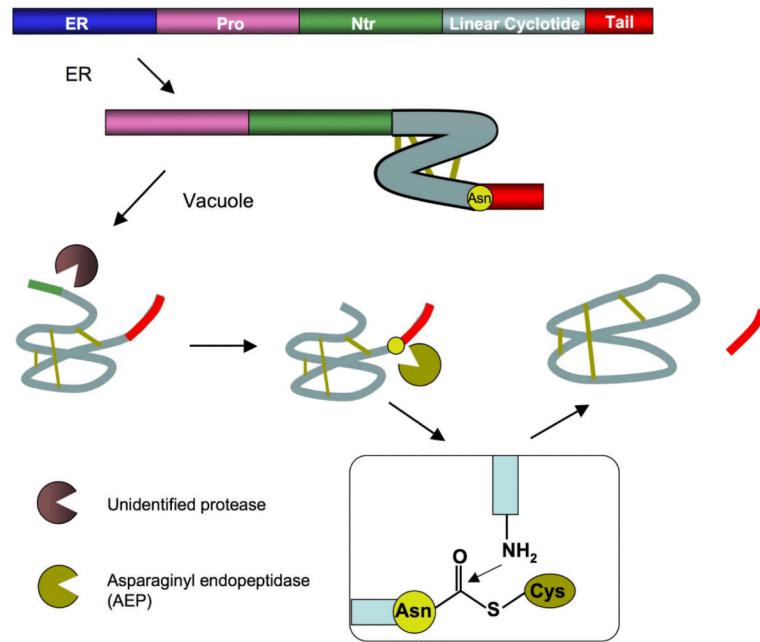
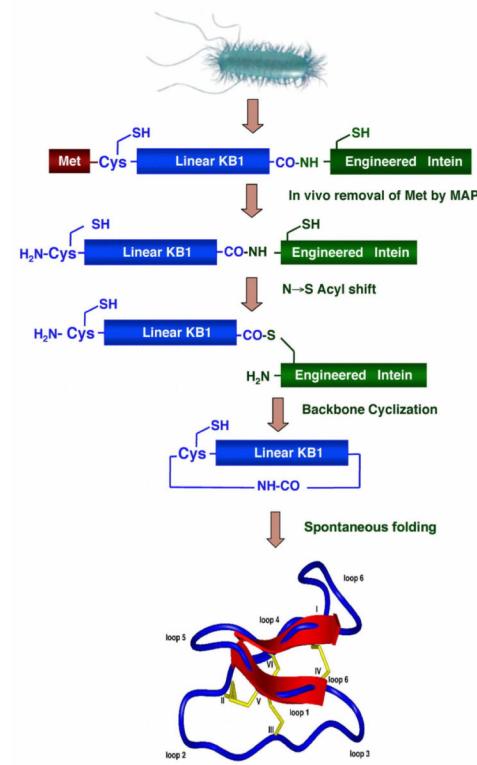


Figure 1.

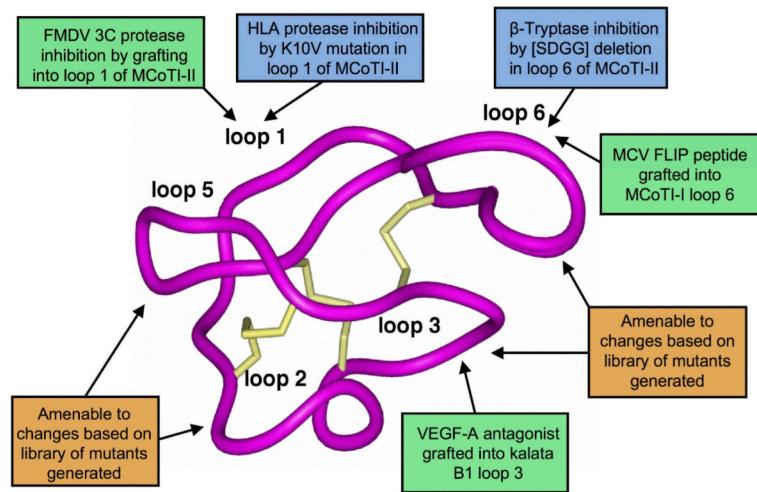
Primary and tertiary structures of representative cyclotides from the bracelet (kalata B1; pdb ID: 1NB1 [108]), Möbius (cycloviolacin O1, pdb ID: 1NBJ [108]), and trypsin inhibitor (MCoTI-II, pdb ID: 1IB9 [109]) subfamilies. Conserved cysteine residues and disulfide bonds are shown in yellow. The blue line denotes the circular backbone.

**Figure 2.**

Schematic representation of the putative mechanism of protease-catalyzed cyclization for cyclotides [28, 29]. The prototypic linear precursor protein (top) comprises an endoplasmic reticulum (ER)-targeting sequence (dark blue), a pro-region (purple), an N-terminal repeat region (Ntr; green), a cyclotide domain (light grey) and a C-terminal tail (red). It also features a conserved asparagine (yellow) at the C-terminal cleavage point of the cyclotide domain. The precursor is processed in the ER and vacuole, disulfide bonds are formed (yellow), and a range of unidentified proteases (brown) trim the precursor. In the final stage, the active-site cysteine of an AEP (yellow) displaces the C-terminal tail to form an enzyme-acyl intermediate (boxed). This intermediate is then attacked by the cyclotide N-terminal glycine to form the mature cyclic peptide. Figure taken from reference [68].

**Figure 3.**

Intein-mediated backbone cyclization for the biosynthesis of cyclotides kalata B1 (kB1) and MCoTI-II in *E. coli* cells [69, 70]. The backbone cyclization of the linear cyclotide precursor is mediated by a modified protein splicing unit or intein. The cyclized product then folds spontaneously in the bacterial cytoplasm.

**Figure 4.**

Using the cyclotide molecular scaffold for drug design. Summary of the changes engineered into the cyclotide framework to introduce new biological functions. Peptide sequences have been successfully grafted into loop 3 of kalata B1 [94] and loops 1 and 6 of MCoTI-II [95] and MCoTI-I (unpublished results), respectively. Cyclotide based libraries have been generated using loops 2, 3, 5 and 6. The cyclotide structure used in the figure corresponds to MCoTI-II (pdb ID: 1IB9 [109]).

Legume cyclotides shed light on the genetic origin of knotted circular proteins

Julio A. Camarero¹

Department of Pharmacology and Pharmaceutical Sciences, School of Pharmacy, University of Southern California, Los Angeles, CA 90033

Cyclotides are fascinating circular proteins ranging from 28 to 37 aa residues that are naturally expressed in plants. They exhibit antimicrobial, insecticidal, antihelmintic, cytotoxic, and antiviral activities (1), and protease inhibitory activity (2), and can exert uterotonic effects (3). They all share a unique head-to-tail circular knotted topology of three disulfide bridges, with one disulfide bond penetrating through a macrocycle formed by the other two disulfides bonds and interconnecting peptide backbones, forming what is called a cystine knot topology (Fig. 1) (1). This cyclic cystine knot framework gives cyclotides a compact, highly rigid structure (4), which confers exceptional resistance to thermal/chemical denaturation and enzymatic degradation (5), thereby making cyclotides a promising molecular scaffold for drug discovery (6, 7). So far, cyclotides have been discovered in plants from the Rubiaceae (coffee), Violaceae (violet), and Cucurbitaceae (squash) families (8, 9), and more recently in the Fabaceae (legume) family (Fig. 1) (10). The discovery of cyclotides in the Fabaceae family of plants represents an important new development because this family of plants is the third largest on Earth, comprising approximately 18,000 different species. Some of these species are widely used as crops in human nutrition and food supply. This opens the intriguing possibility of using these plants for the large-scale production of cyclotides with pharmaceutical or agrochemical properties by using transgenic crops. The key to accomplishing that, however, is to have a better understanding of the mechanism that produces these interesting microproteins in this family of plants.

The report by Poth et al. in PNAS (11) brings us closer to that exciting possibility by describing the gene encoding the protein precursor of a unique cyclotide (Cter M) isolated from the leaf of butterfly pea (*Clitoria ternatea*), a representative member of the Fabaceae plant family. All the cyclotides reported so far from the Violaceae and Rubiaceae families are biosynthesized via processing from dedicated genes that, in some cases, encode multiple copies of the same cyclotide, and in others, mixtures of different cyclotide sequences (Fig. 1) (12). Poth et al. (11) reveal that the sequence encoding the cyclotide Cter M, however, is embedded within the albumin-1 gene of *C. ternatea*

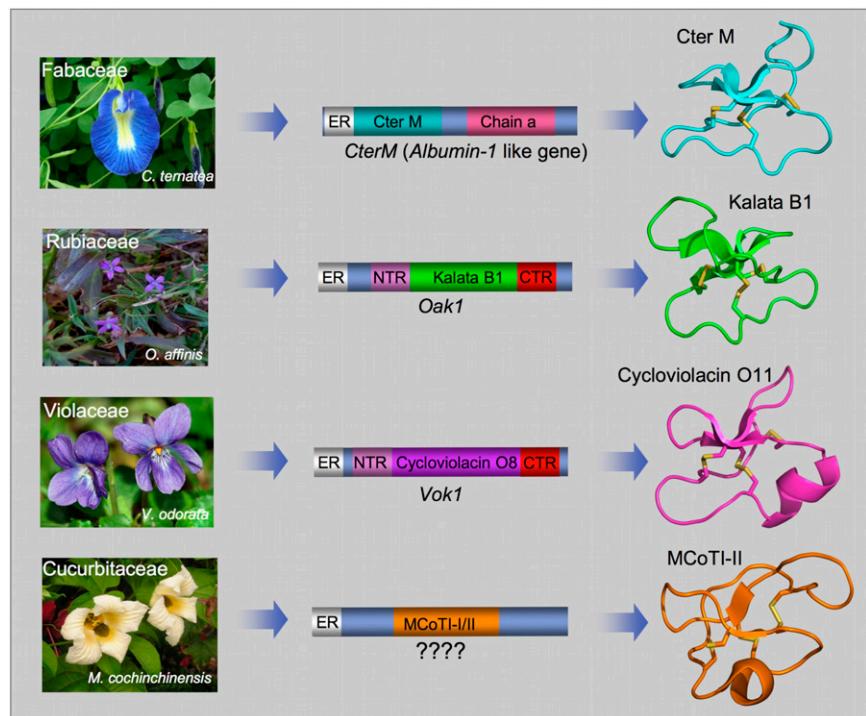


Fig. 1. Genetic origin of cyclotides from different plant families. Rubiaceae (*Oldenlandia affinis*) and Violaceae (*Viola odorata*) plants have dedicated genes for the production of cyclotides (12). These cyclotide precursors comprise an ER signal peptide, an N-terminal Pro region, the N-terminal repeat (NTR), the mature cyclotide domain, and a C-terminal flanking region (CTR). In contrast, the *CterM* gene (*C. ternatea*, Fabaceae) shows an ER signal peptide immediately followed by the cyclotide domain, which is flanked at the C terminus by a peptide linker and the albumin a-chain. The *CterM* cyclotide domain replaces albumin-1 b-chain. The genetic origin of the Cucurbitaceae cyclotides (found in the seeds of *M. cochinchinensis*) remains to be identified.

(Fig. 1). Plant albumins are part of the nutrient reservoir, but they also play a role in host defense. Generic albumin-1 genes are comprised of an ER signal sequence followed by an albumin chain-b, a linker, and an albumin chain-a. In the precursor of cyclotide Cter M, the cyclotide domain replaces the albumin chain-b domain. This interesting finding raises the question of how this replacement took place in evolution. There are two possibilities: (i) gradual evolution of the chain-b domain into the cyclotide domain or (ii) rapid lateral transfer of the cyclotide gene into the albumin gene. Poth et al. (11) present evidence supporting a gradual evolutionary path, whereby the albumin-1 chain-b slowly evolved into a more stable cyclotide domain. For example, the pea albumin-1 subunit-b (PA1b), one of the best-studied Fabaceae albumin components, is a 37-aa peptide from pea seeds (*Pisum sativum*), which also contains a cystine-knot structure

(13). Remarkably, the cystine-knot core of PA1b overlays extremely well with that of the cyclotide Cter M. The composition and size of the PA1b loops are, however, totally different from those of the cyclotide Cter M. Recent mutagenesis studies on PA1b have also recently shown that this albumin domain is highly tolerant to mutations outside the cystine knot core (14). These observations support the possibility of divergent evolution of cyclotides from ancestral albumin domains, wherein evolution and natural selection provided an alternative loop decoration of the original cystine knot albumin core.

A final question remains. Linear cystine knot proteins such as PA1b are not backbone

Author contributions: J.A.C. wrote the paper.

The author declares no conflict of interest.

See companion article on page 10127.

¹E-mail: jcamarer@usc.edu.

cyclized like cyclotides. What made the cyclization process possible, allowing the final transformation of an evolved cystine-knot albumin domain into a cyclotide? As indicated by Poth et al. (11), the structural analysis of PA1b may reveal some clues about how this could have happened. The NMR structure of PA1b (13) reveals that the N and C termini are very close to each other, and it is possible that mutations in the albumin genes predisposed them to cyclization during the evolution process.

However, what type of mutations could allow the backbone cyclization of a linear cystine knot albumin domain? Although the complete mechanism of how cyclotide precursors are processed and cyclized has not been fully characterized yet, recent studies indicate that an asparaginyl endopeptidase (AEP; also known as vacuolar processing enzyme or legumain) is a key element in the cyclization of cyclotides (15, 16). It has been proposed that the cyclization step mediated by AEP takes place at the same time as the cleavage of the C-terminal propeptide from the cyclotide precursor protein through a transpeptidation reaction (15). The transpeptidation reaction involves an acyl-transfer step from the acyl-AEP intermediate to the N-terminal residue of the cyclotide domain (16). A similar process has been used for the chemical (17), chemoenzymatic (18), and recombinant (19) production of cyclotides. AEPs are Cys proteases that are very common in plants and are able to specifically cleave the peptide bond at the C terminus of Asn and, less efficiently, Asp residues. All the cyclotide precursors identified so far, including those from *C. ternata*, contain a well conserved Asn/Asp residue at the C terminus of the cyclotide domain, which is consistent with the idea that cyclotides are cyclized by a transpeptidation reaction mediated by AEP (15).

Despite these similarities, *C. ternata* cyclotides also show some differences regarding the residue immediately following the mechanistically conserved Asn. In the cyclotide precursors from the Violaceae and

Rubiaceae families, the C-terminal Asn/Asp residue is always followed by a small amino acid, either Gly or Ser. However, the Cter M precursor reported by Poth et al. (11) indicates that a small amino acid is not always required in that position. Moreover, some *C. ternata* cyclotides also have a His residue at the N terminus of the cyclotide precursor rather than the most common Gly residue found in most cyclotide domains (10). These observations seem to indicate that, at least in the Fabaceae family, the AEP-mediated transpeptidation step may be more tolerant than previously recognized.

The finding that albumin genes can evolve into protein precursors that can be subsequently processed to become cyclic was described in a recent report on the biosynthesis of the sunflower trypsin inhibitor peptide, SFTI-1 (20). SFTI-1 is a 14-residue peptide isolated from sunflower seeds with a head-to-tail cyclic backbone structure having only a single disulfide bond. In this case, the SFTI-1 linear precursor is embedded within a “napin-type” 2S albumin.

The report by Poth et al. (11) indicates that the biosynthetic origin of some cyclotides are very different from others, which could suggest that cyclic peptides might be more widely distributed than is currently realized. The exceptional stability of backbone-cyclized peptides may give them an evolutionary advantage, which may provide the driving force for the evolution of multiple biosynthetic pathways including the use of dedicated or recycled genes, with albumins now being implicated in the biosynthesis of two different classes of cyclic peptides.

In this context, it is worth noting that the protein precursors of the only two cyclotides isolated so far from the Cucurbitaceae plant family, *Momordica cochinchinensis* trypsin inhibitor I and II (MCoTI-I/II; Fig. 1), remain yet to be identified. These cyclotides are found in the seeds of *M. cochinchinensis* (a tropical squash plant) and are potent trypsin inhibitors. MCoTI cyclotides do not share significant sequence homology with the

other cyclotides beyond the presence of the three-cystine bridges that adopt a similar backbone-cyclic cystine-knot topology (Fig. 1) and are more related to linear cystine-knot squash trypsin inhibitors. In fact, an acyclic version of MCoTI-cyclotides (known as MCoTI-III) can also be found in the seeds of *M. cochinchinensis*. This situation, in which the cyclic and linear versions of the cys-knot protein coexist in the same organism, provides a unique opportunity to study the genetic origin and evolution of these interesting molecules. Identification of the protein precursors for the cyclic and linear versions of these cystine-knot trypsin inhibitors should provide a unique snapshot in the evolutionary process of plant cyclic cystine-knot proteins.

In summary, the work by Poth et al. (11) provides critical information on the origin, evolution, and processing of cyclotides from a plant of the Fabaceae family. The discovery of unique cyclotides as well as other cyclic peptides from a wide range of plants is key to define and fully understand the different cyclization mechanisms used by plants. So far, the expression of cyclotides in transgenic plants has been attempted only in *Arabidopsis* and tobacco (15, 16), in which cyclotide expression is highly inefficient, giving rise to mostly acyclic or truncated proteins. The proven ability of *C. ternata* to produce fully folded cyclotides seems to suggest that other species of the Fabaceae family could also be used for the production of cyclotides. Several members of this large family of plants are agricultural crops, which opens the intriguing possibility of generating genetically engineered crops for the large-scale production of cyclotides with useful pharmacological or agrochemical properties in the near future.

ACKNOWLEDGMENTS. This work was supported by National Institutes of Health Research Grant R01-GM090323 and Department of Defense Congressionally Directed Medical Research Program Grant PC09305.

1. Daly NL, Rosengren KJ, Craik DJ (2009) Discovery, structure and biological activities of cyclotides. *Adv Drug Deliv Rev* 61:918–930.
2. Avrutina O, et al. (2005) Trypsin inhibition by macrocyclic and open-chain variants of the squash inhibitor MCoTI-II. *Biol Chem* 386:1301–1306.
3. Gran L, Sandberg F, Sletten K (2000) Oldenlandia affinis (R&S) DC. A plant containing uteroactive peptides used in African traditional medicine. *J Ethnopharmacol* 70: 197–203.
4. Puttamadappa SS, Jagadish K, Shekhtman A, Camarero JA (2010) Backbone dynamics of cyclotide MCoTI-I free and complexed with trypsin. *Angew Chem Int Ed Engl* 49:7030–7034.
5. Colgrave ML, Craik DJ (2004) Thermal, chemical, and enzymatic stability of the cyclotide kalata B1: The importance of the cyclic cystine knot. *Biochemistry* 43: 5965–5975.
6. Jagadish K, Camarero JA (2010) Cyclotides, a promising molecular scaffold for peptide-based therapeutics. *Bio-polymers* 94:611–616.
7. Garcia AE, Camarero JA (2010) Biological activities of natural and engineered cyclotides, a novel molecular scaffold for peptide-based therapeutics. *Curr Mol Pharmacol* 3:153–163.
8. Gruber CW, et al. (2008) Distribution and evolution of circular miniproteins in flowering plants. *Plant Cell* 20: 2471–2483.
9. Chiche L, et al. (2004) Squash inhibitors: From structural motifs to macrocyclic knottins. *Curr Protein Pept Sci* 5:341–349.
10. Poth AG, et al. (2011) Discovery of cyclotides in the Fabaceae plant family provides new insights into the cyclization, evolution, and distribution of circular proteins. *ACS Chem Biol* 6:345–355.
11. Poth AG, Colgrave ML, Lyons RE, Daly NL, Craik DJ (2011) Discovery of an unusual biosynthetic origin for circular proteins in legumes. *Proc Natl Acad Sci USA* 108: 10127–10132.
12. Dutton JL, et al. (2004) Conserved structural and sequence elements implicated in the processing of gene-encoded circular proteins. *J Biol Chem* 279:46858–46867.
13. Jouvensal L, et al. (2003) PA1b, an insecticidal protein extracted from pea seeds (*Pisum sativum*): 1H-2D NMR study and molecular modeling. *Biochemistry* 42:11915–11923.
14. Da Silva P, et al. (2010) Molecular requirements for the insecticidal activity of the plant peptide pea albumin 1 subunit b (PA1b). *J Biol Chem* 285:32689–32694.
15. Saska I, et al. (2007) An asparaginyl endopeptidase mediates *in vivo* protein backbone cyclization. *J Biol Chem* 282:29721–29728.
16. Gillon AD, et al. (2008) Biosynthesis of circular proteins in plants. *Plant J* 53:505–515.
17. Craik DJ, Conibear AC (2011) The chemistry of cyclotides. *J Org Chem*, in press.
18. Thongyoo P, Roqué-Rosell N, Leatherbarrow RJ, Tate EW (2008) Chemical and biomimetic total syntheses of natural and engineered MCoTI cyclotides. *Org Biomol Chem* 6:1462–1470.
19. Austin J, Wang W, Puttamadappa S, Shekhtman A, Camarero JA (2009) Biosynthesis and biological screening of a genetically encoded library based on the cyclotide MCoTI-I. *ChemBioChem* 10:2663–2670.
20. Mylne JS, et al. (2011) Albumins and their processing machinery are hijacked for cyclic peptides in sunflower. *Nat Chem Biol* 7:257–259.